Eine Einführung in die Erkenntnistheorie

Vorlesung von Wolfgang Spohn im SS 1995

Skriptum ausgearbeitet von Martin Rechenauer

Inhaltsverzeichnis

1. Zur Einführung

Die Erkenntnistheorie ist ohne Zweifel eine der zentralen Grunddisziplinen der Philosophie; diese Position nimmt sie freilich erst seit Descartes, also seit der philosophischen Neuzeit, ein.

Es ist kein schlechtes Verfahren, die Gebiete der Philosophie anhand der Leitfragen einzuteilen, deren Beantwortung seit der Antike versucht wird. Da beginnt man mit der Leitfrage der *prima philosophia*, welche nach klassischem Verständnis in der Metaphysik und Ontologie besteht: Was gibt es? Die Leitfrage der Erkenntnistheorie hingegen lautet: Was kann ich wissen? In der Neuzeit wurde die klassische ontologische Fragestellung dann zusehends erkenntnistheoretischen Gesichtspunkten unterworfen. Mit Descartes kam die Idee auf, daß man zunächst einmal klären müsse, was man wissen könne, bevor man sich der Frage widmen könnte, was es gibt. Die Entwicklung der modernen analytischen Philosophie setzte sich dahin fort, daß nunmehr die Sprachphilosophie eine Schlüsselrolle einnimmt. Deren Leitfrage wiederum lautet: Wie reden wir über das, was es möglicherweise gibt und wovon wir etwas wissen können? Allerdings ist die sprachphilosophische Wende in der modernen analytischen Philosophie ebenfalls erkenntnistheoretisch motiviert; da Wissensansprüche allgemein in sprachlicher Gestalt einhergehen, liegt es nahe, sie in dieser ihrer typischen Erscheinungsform zu untersuchen. Somit bleibt die Erkenntnistheorie mit ihrer spezifischen Fragestellung weiterhin im Zentrum der Aufmerksamkeit.

Es ist bei einer Vorlesung der vorliegenden Art meist üblich, einen historischen Abriß zu präsentieren, in dessen Verlauf dann die Probleme samt der Art und Weise, wie sie von den Altmeistern der Philosophie behandelt wurden, vorgestellt werden. In dieser Vorlesung soll jedoch ein systematischer Zugang vorherrschen; das hat mit den Vorlieben und Talenten dessen, der sie hält, zu tun, wie auch mit seiner Überzeugung, daß die klassischen Fragen sich in der aktuellen Diskussion oft genug wiederfinden und dort häufig klarer behandelt werden.

Ich habe oben gesagt, daß die Leitfrage der Erkenntnistheorie sei: Was kann ich wissen? Nun ist der Wissensbegriff selbst Gegenstand erkenntnistheoretischer Analysen, die recht vielfältig ausfallen können, aber stets als Kern die Behauptung aufweisen, daß Wissen wahre Überzeugung oder wahrer Glauben sei.[1] Damit erweist sich aber der Glaubensbegriff, oder wie er vielleicht sprachlich eindeutiger bezeichnet werden sollte: der Begriff der Überzeugung als der eigentliche Grundbegriff. Entsprechend lautet die Grundfrage denn auch angemessener: Was soll ich glauben?

Noch präziser kann man die Grundbegriff der Erkenntnistheorie mit Hilfe der folgenden Formulierung zum Ausdruck bringen:

Subjekt a glaubt zum Zeitpunkt t im Grade r, daß p.

Hierin haben wir vier Positionen, die schematischer Natur sind. Zu den ersten beiden, zum Subjekt der Erkenntnis a und zu den Zeitpunkten t, wird im folgenden nichts mehr gesagt werden; sie sind keine spezifischen Gegenstände der Erkenntnistheorie. Die für unsere Disziplin relevanten Parameter sind dagegen der Überzeugungsgrad r und der Überzeugungsgrad p; sie werden in den zentralen mittleren Kapiteln dieser Vorlesung genauer behandelt.

Auf zwei Dinge möchte ich einleitend noch hinweisen. Erstens wurde bei der obigen Reformulierung der Leitfrage der Erkenntnistheorie eine subtile Veränderung vorgenommen: aus „Was kann ich wissen?" wurde „Was soll ich glauben?". Die neue Formulierung enthält ein normatives Element, das nach einer näheren Erklärung ruft. Diese Erklärung ergibt sich im Zuge der Beantwortung einer zweiten Frage, auf die man angesichts der Formulierung der erkenntnistheoretischen Grundsituation kommen könnte: Was unterscheidet die Untersuchung des Erkenntnistheoretikers, also eines Philosophen, eigentlich von einer Untersuchung im Rahmen der kognitiven Psychologie? Auch dort ist man bemüht, die Überzeugungen von Subjekten zu analysieren und besser zu verstehen. Hinzu kommt, daß philosophische und psychologische Untersuchungen in diesen Bereichen ein gemeinsames Erbe aufweisen und die Disziplinen erst in neuerer Zeit ihre eigenen Wege einschlugen.

[1] Terminologische Anmerkung: Die Substantivierung des Verbs „glauben" im Deutschen ist hochgradig mißverständlich, nicht zuletzt wegen der religiösen Nebenbedeutung. Nun gibt es im Deutschen die harmlose Verwendung „Hans glaubt, daß die Sonne scheint", welche ausdrückt, daß Hans eine Überzeugung hat, die den Inhalt aufweist, daß die Sonne scheint. Das führt schon auf das nächste Problem: es gilt zu erklären, was Inhalte sind. Darauf wird im Kapitel 7 weiter eingegangen werden. Im folgenden werden wir „Überzeugung" als die Substantivierung von „glauben" verwenden. Man könnte auch „Meinung" verwenden, aber das hat ebenfalls unerwünschte Nebenbedeutungen. Im Englischen ist es einfacher: da gibt es das Verb „to believe", dessen Substantivierung „belief" lautet. Das Substantiv weist keine religiösen Konnotationen auf, die ihren Ausdruck vielmehr in dem Substantiv „faith" finden. Allerdings gibt es beim Verb eine analoge Doppeldeutigkeit wie im Deutschen: den propositionalen Gebrauch „glauben, daß die Sonne scheint", „to believe that the sun is shining" gegenüber „an Gott glauben", „to believe in god". In unserer Diskussion wird nur der propositionale Sinn von „glauben" eine Rolle spielen.

Doch es gibt drei wesentliche Unterschiede. (1) Psychologen sind als empirisch arbeitende Wissenschaftler an empirisch fruchtbaren Anbindungen ihrer Fragestellung interessiert, wohingegen Philosophen begriffliche Fragen und Explikationen in den Mittelpunkt ihrer Bemühungen zu stellen pflegen. (2) Die Glaubenstheorie in der Form, wie sie in der Erkenntnistheorie getrieben wird, ist keine rein empirische, sondern eine wenigstens teilweise normative Disziplin – wie es in der Reformulierung der Leitfrage ja zum Ausdruck kam: „Was *soll* ich glauben?“ Das heißt, Philosophen sind eher daran interessiert, zu verstehen, was man vernünftigerweise in einer Welt, die so oder so beschaffen ist, glauben sollte. Etwas überspitzt gesagt: Philosophen streben nach einer normativen Theorie, die erklärt, was es heißt, in bestimmten Situationen rationalerweise etwas zu glauben oder rational seine Überzeugungen zu ändern, während Psychologen sich eher dafür interessieren, wie die Leute faktisch ihre Überzeugungen erwerben und verändern. (3) Philosophen sind vor allem an der Wahrheitsfrage interessiert. Können wir herausfinden, welche der Überzeugungen eines Subjektes wahr sind und welche nicht – und wenn ja, wie? Deswegen stand ja traditionell der Wissens- und nicht der Glaubensbegriff im Zentrum der Erkenntnistheorie. Psychologen werden sich damit jedoch nicht ausgiebig herumschlagen; ihnen ist es eher um die faktischen Überzeugungen der Menschen zu tun, gleichgültig, ob sie wahr sind oder nicht. Berücksichtigt man alle diese Überlegungen, sollte klar sein, daß die folgenden Überlegungen nicht psychologischer Natur sein werden. Das schließt allerdings überhaupt nicht aus, daß es vielfältige Möglichkeiten zur wechselseitigen Befruchtung zwischen Philosophie und empirischer Psychologie gibt.

Die Vorlesung hat insgesamt den folgenden Aufbau: Im Kapitel 2 wird anhand historischen Materials die skeptizistische Herausforderung behandelt, die darin liegt, daß es auf die Frage „Können wir überhaupt etwas wissen?“, angeblich keine überzeugende positive Antwort gibt, und die traditionell im Zentrum erkenntnistheoretischer Diskussionen steht; dies soll gleichzeitig dazu motivieren, zu den späteren theoretischeren Kapiteln fortzuschreiten. Die Skeptizismus-Diskussion führt die Notwendigkeit vor Augen, zunächst einmal den Wissensbegriff genauer zu klären. Dies ist Gegenstand des Kapitels 3, in dem verschiedene Theorien über den Begriff des Wissens, die in der analytischen Philosophie im Schwange sind, vorgestellt werden. Diese zeigen, wie schon angedeutet, daß der Begriff des Glaubens eigentlich der grundlegende Begriff der Erkenntnistheorie ist, um den es dann in den Kapiteln 4-7 geht. Die Kapitel 4-6 sind zunächst den Graden des Überzeugtseins gewidmet; das klingt so dürr, ist aber in Wahrheit zentral. Genauer gesagt, befaßt sich das Kapitel 4 mit der statischen Glaubenstheorie, worin die grundlegenden Möglichkeiten zur theoretischen Erfassung des Glaubensbegriffs vorgestellt werden. Das Kapitel 5 beschäftigt sich dann mit den dynamischen Aspekten der Glaubenstheorie, durch welche sie erst vollständig wird und die in der Tat zentral mit dem Induktionsproblem verknüpft sind. Es besteht ferner ein enger Zusammenhang zwischen der Dynamik von Überzeugungen einerseits und dem philosophisch wichtigen Begriff des Grundes oder der Begründung andererseits; dies ist Thema des Kapitels 6. Im Kapitel 7 geht es um den Parameter *p* des obigen vierstelligen Grundbegriffs. Genauer gesagt, geht es sowohl um Glaubenszuschreibungen als auch um Glaubensinhalte oder um die Gegenstände des Glaubens, also das, wofür *p* stellvertretend steht; die Hauptschwierigkeit dabei wird sein zu erkennen, daß beide Themen sehr eng miteinander zusammenhängen, ohne daß man sie miteinander identifizieren darf. Ging es so weit nur ums Glauben selbst und nichts weiter, so sind die Kapitel 8 und 9 den kausalen Verbindungen gewidmet, in denen unsere kognitiven Zustände stehen; die Einsicht in die Beschaffenheit dieser Verbindungen befördert unser Verständnis dieser Zustände ganz wesentlich. Da geht es einmal, im Kapitel 8, darum, wie sich Überzeugungen in Handlungen und speziell in Äußerungen manifestieren. Zum andern sollte es im Kapitel 9 um die philosophische Wahrnehmungstheorie gehen, die genauer das Verhältnis zwischen Wahrnehmungen und Überzeugungen analysiert. In der Vorlesung fehlte dafür freilich die Zeit, so daß ich mich hier hur auf einige Verweise beschränke. Insoweit ist dieses Skriptum also ein Torso. Das letzte Kapitel 10 kehrt wieder zur Wahrheitsfrage zurück, indem es den Zusammenhang von Erkenntnistheorie und Metaphysik näher betrachtet.[2]

---

[2] Einen allgemeinen Überblick über Fragestellung der Erkenntnistheorie erhält man z.B. aus dem Sammelband von Bieri (1987). Dieser Band enthält wichtige Beiträge aus der analytischen Diskussion um den Wissensbegriff; besonders empfehlenswert sind die allgemeine Einleitung und die Einleitungen zu den Unterabschnitten von Bieri selber. Einen fortgeschritteneren Überblick zum Thema gibt von Kutschera (1982).

2. Die skeptizistische Herausforderung

Am Beginn steht wie so oft eine recht trivial anmutende Beobachtung. Jeder Mensch hat eine Menge Überzeugungen höchst unterschiedlichen Inhalts. So glaube ich z.B., daß die Bundesrepublik Deutschland 1949 gegründet wurde, daß es moralisch verwerflich ist, Menschen zu foltern, daß 2+2 = 4 und vieles andere mehr. Wir werden uns aus Gründen der Einfachheit aber im folgenden nur auf solche Überzeugungen konzentrieren, die auf unkontroverse Weise als wahr oder falsch aufgefaßt werden können. Überzeugungen über moralische Fragen haben sicher nicht diesen unkontroversen Status; mathematische Überzeugungen werfen auch Probleme eigener Art auf, mit denen wir uns im Moment nicht herumschlagen wollen. Also möge es weiterhin nur um Überzeugungen gehen, die Annahmen über die empirische Beschaffenheit der Welt zum Ausdruck bringen.

Warum diese Fixierung auf Überzeugungen? Überzeugungen sind wichtig für Wesen der Art, wie wir es sind. Wir sind Wesen mit Interessen, Zielen und Absichten. Unsere Überzeugungen über die Beschaffenheit der Welt sind wichtige Vermittlungsfaktoren, wenn wir uns in dieser Welt betätigen und unsere Ziele erreichen wollen. Die Annahme scheint plausibel, daß wir unsere Ziele desto besser in die Tat umsetzen können, je mehr unsere Überzeugungen über die Welt der Wahrheit entsprechen. Das betrifft sowohl die Wahrheit unserer Überzeugungen über die Umstände in der Welt, unter denen wir handeln wollen, als auch die über die Zweck-Mittel-Relationen, die wir veranschlagen.

In der Tat ist eine Überzeugung etwas, was wesentlich entweder wahr oder falsch ist. Das zeigt sich schon daran, daß man statt „ich glaube, daß *p*“ ebenso gut sagen kann: „ich halte *p* für wahr“; allenfalls bei krassem und völlig irrationalen Wunschdenken könnte man vielleicht davon reden, daß beides auseinanderfällt. Generell kann man folgendes sagen: Die Überzeugung eines Subjekts *a*, daß *p*, ist genau dann wahr, wenn „*p*“ wahr ist (d.h. wenn es der Fall ist, daß *p*). Der Inhalt der Überzeugung ist mithin zugleich ihre Wahrheitsbedingung. An dieser Stelle sollte man aber gleich auch auf einen wesentlichen Unterschied zwischen Wissen und Glauben hinweisen: Wissen *muß* wahr sein, Überzeugungen müssen nur wahr oder falsch sein können.

Der andere wichtige Aspekt von Überzeugungen neben ihrem Bezug auf Wahrheit ist ihr Verhältnis zu guten Gründen. Wenn ich glaube, daß *p*, dann habe ich, zumindest wenn ich ein einigermaßen rationales Wesen bin, Gründe für diese Überzeugung. Finde ich in mir eine Überzeugung, die offenbar nicht durch Gründe gestützt ist, so ist das ein hinreichender Grund, sie mit Vorsicht zu betrachten und sie sogar aufzugeben, falls sich nach genauerer Untersuchung keine Gründe für sie finden lassen.

Überzeugungen können also wahr oder falsch sein und sind durch mehr oder weniger gute Gründe gestützt. Nun aber tritt der Skeptizist auf. Wie die bisherigen Ausführungen schon nahelegen, kann er das in zwei Gestalten tun: in der einen richtet er sich gegen den Wahrheitsanspruch, der mit unseren Überzeugungen einhergeht, in der anderen greift er den Anspruch auf Begründetheit an. Zwei herausragende Gestalten der Philosophiegeschichte sollen im folgenden als Protagonisten der jeweiligen Varianten dienen: René Descartes und David Hume.[3] Dabei ist zu beachten, daß Descartes selbst gerade kein Skeptizist ist; er führt in seiner bekannten Argumentation in den *Meditationen* die Figur des Skeptizisten nur ein, um im Zuge der Widerlegung von dessen Position ein sicheres Fundament für Wissensansprüche zu ermitteln.

Der Skeptizist beginnt ganz harmlos. Er stellt nur Fragen, die darauf abzielen, daß wir unsere Wahrheits- und Begründungsansprüche besser verteidigen und auf eine sicherere Grundlage stellen. Er fragt z.B., woher wir wissen, daß die Linie 21 zum Jahnplatz fährt oder daß die Bundesrepublik 1949 gegründet wurde. Doch welche Antworten er auch bekommt, er hört nicht auf mit dem Fragen. Dabei werden die Fragen immer seltsamer, wie man etwa an der folgenden sieht: Woher wissen wir, daß die Welt nicht eben vor 5 Minuten zu existieren begonnen hat? Gewiß, wir scheinen alle zu glauben, daß wir erheblich älter sind als 5 Minuten – unsere Erinnerungen bestätigen das jederzeit. Aber wenn diese Erinnerungsspuren, die wir alle in uns tragen, vor 5 Minuten ebenfalls zu existieren begonnen haben, so daß es uns nur so scheint, als hätten wir schon sehr viel länger gelebt? Wie können wir so einen Fall ausschließen? Angesichts dieses Trommelfeuers von Fragen entsteht der Verdacht, daß der Skeptizist ein unfaires Spiel spielt. Es scheint gar nicht möglich, ihm eine Antwort zu geben, die ihn befriedigt und die hilft, seine Fragerei zu beenden; er wird immer wieder einen neuen Ansatzpunkt finden.

An dieser Stelle ergeben sich mehrere Reaktionsmöglichkeiten. Erstens könnte man in einen Zustand der Urteilslosigkeit und des endlosen Zweifels verfallen. So

[3] Die klassischen Texte hier sind Descartes (1641) und Hume (1739).

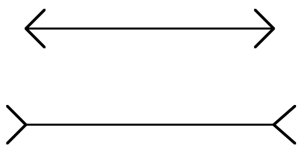
soll der antike Skeptiker Pyrrho sich tatsächlich jeglichen Urteils enthalten haben und auf diese Weise doch ein ganz zufriedener Mensch gewesen sein. Allerdings sind auch nachgerade psychotische Zweifelsreaktionen denkbar, oder der Umschlag des unausgesetzten Zweifels in merkwürdige positive Thesen. So ist in der Philosophie der Solipsismus, also die These, daß nur die eigenen individuellen Bewußtseinszustände existieren, oft als eine Konsequenz aus dem Skeptizismus angesehen worden. Die „normale" Reaktion dürfte aber die Einstellung sein, die auch der Skeptizist Hume selbst eingenommen hat, der bekennt, in seinem Arbeitszimmer ein Skeptizist zu sein, aber dann durch die Ablenkungen des Alltags von den skeptischen Zweifeln befreit wird. Er kann die Herausforderung zwar nicht beantworten, aber erkennt doch, daß sie für ihn keine praktische Bedeutung hat (cf. die bekannte Stelle in Hume 1739, I.IV.VII).

Doch geben sich nicht alle mit dieser Einstellung zufrieden. So besteht eine zweite Reaktion in dem Versuch, im skeptizistischen Katz-und-Maus-Spiel doch irgendwelche falsche Annahmen zu finden, die dieses Spiel so schief laufen lassen. So könnte es sich etwa herausstellen, daß wir mit einem zu starken Begriff von Wissen, Wahrheit oder Begründung begonnen haben und der Skeptizist genau dies für seine Fragerei ausnützt.

Die dritte mögliche Reaktion besteht in der Einnahme eines Standpunkts jenseits des Frage-und-Antwort-Spiels, indem man versucht, dem Skeptiker eine Art Regelverstoß nachzuweisen. So könnten unsere Kommunikationsregeln etwa nur dem echten Zweifler ein Recht auf Beantwortung seiner Fragen geben, der Skeptiker sich jedoch durch seine Art der Fragerei als echter Zweifler disqualifizieren. Oder es könnte sein, daß man wohl jede einzelne Überzeugung, jede einzelne These anzweifeln darf, daraus aber nicht folgt, daß man *alle zusammen* in Zweifel ziehen kann; der unzulässige Schritt des Skeptikers besteht vielleicht genau darin. Es lohnt sich, diese dritte Reaktion philosophisch zu vertiefen. Hier werde ich vor allem die zweite Reaktionsmöglichkeit ausführen.

Um mit dem Fall eines skeptizistischen Angriffs auf den Wahrheitsanspruch von Erkenntnissen zu beginnen, bietet sich Descartes' methodischer Zweifel als Illustration an. Wie schon betont, ist Descartes kein Skeptiker – im Gegenteil, er treibt den skeptischen Zweifel auf die Spitze, um ihn dann zu überwinden und Wissen auf eine sichere Grundlage zu stellen. Ich möchte vorsichtshalber noch darauf hinweisen, daß die folgenden Ausführungen nicht den Anspruch erheben können, eine schulmäßige Descartes-Exegese abzugeben. Wesentliche Elemente von Descartes' Philosophie, etwa die zentrale Unterscheidung von *res cogitans* und *res extensa* bleiben völlig außer acht. Und die Subtilitäten der Argumentation lassen sich in der gebotenen Kürze auch nicht angemessen berücksichtigen.

Der Ausgangspunkt von Descartes' Gedankenexperiment aus der ersten der *Meditationen* ist wieder eine vertraute Beobachtung. Informationen sind häufig unzuverlässig; wir täuschen uns oft. Das kann passieren, weil wir uns auf unzuverlässige Informanten verlassen, weil wir aus eigenen Erfahrungen falsche Schlußfolgerungen ziehen, weil wir Sinnestäuschungen unterliegen können, oder sonst irgendetwas schief läuft. Das klassische Beispiel einer Differenz von Sein und Schein ist der ins Wasser gehaltene gerade Stock, der geknickt erscheint. Wohl ebenso bekannt ist die Müller-Lyer-Täuschung:



Die Strecke zwischen den Pfeilen ist beidemale gleich lang, auch wenn es nicht so scheint. Viele andere Beispiele von möglichen oder tatsächlichen Irrtümern fallen einem schnell ein. Da lesen und hören wir derzeit von Pipeline-Unglücken in Sibirien, wo sich infolgedessen riesige Ölseen in der Taiga ausbreiten sollen. Stimmt das wirklich, ist es nicht nur eine kräftige Übertreibung eines Sensationsjournalismus? Ein weiteres nettes Beispiel ist die alte Mär, daß im Spinat besonders viel Eisen enthalten sein soll. Jahrzehntelang wurden Kinder mit allen Mitteln angehalten, etwas zu essen, was viele von ihnen nicht mochten – nur weil ihre Eltern von der besonderen Qualität dieses Gemüses überzeugt waren. Tatsächlich hat sich in eine bekannte Darstellung des Nährwertes von Spinat, erstellt Ende des 19. Jahrhunderts, ein schlichter Dezimalkomma-Fehler eingeschlichen, der dann in allen Lehrbüchern für Hauswirtschaft über Generationen abgeschrieben wurde.[4] Es gibt noch drastischere Beispiele. Manche werden das Gefühl kennen, wie es ist, wenn man nicht weiß, ob man träumt oder schon wach ist. Leute im Drogenrausch oder auf Entzug haben offenbar die merkwürdigsten, zweifelhaftesten Erfahrungszustände. Alle diese Fälle illustrieren eins deutlich: es könnte auf zuweilen dramatische Art anders sein, als es aussieht – vielleicht ist ja wirklich alles ganz anders.

[4] Falls das auch nicht stimmt, ist es wenigstens hübsch erfunden.

Bei Descartes tritt nun ein Betrügergott auf, der seine Freude daran hat, die armen Subjekte zu verwirren (1. Meditation, Abschnitt 12). Eine hübsche modernere Variante davon sind die von Hilary Putnam erfundenen Hirne im Topf.[5] Stellen Sie sich vor, die Welt ist ganz anders aus als wir es normalerweise annehmen: statt Bäumen, Universitäten, Menschen aus Fleisch und Blut etc. gibt es nur einen riesigen Supercomputer und einen Topf mit Nährlösung, in dem Hirne vor sich hin dümpeln. Die afferenten und efferenten Nervenbahnen dieser Hirne sind an den Computer angeschlossen, der diesen Hirnen die Existenz einer realen Außenwelt vorspiegelt. Diese armen Hirne glauben also z.B., daß sie Studenten an der Uni Bielefeld sind, wo sie eine Vorlesung über Erkenntnistheorie hören, in welcher der Professor so einen Unsinn erzählt, daß Philosophen die Möglichkeit diskutieren, daß es Hirne im Topf geben könnte, die.... Und so weiter. Sind wir vielleicht Hirne im Topf? Woher wissen wir, daß wir keine sind?[6] Putnams Zweifelsexperiment erweist sich freilich als nicht ganz so radikal wie dasjenige von Descartes; dessen Betrügergott spiegelt uns nicht nur die Existenz einer – im skeptizistischen Kontext nicht vorhandenen – Außenwelt vor, er soll uns sogar auch noch über mathematische Wahrheiten täuschen können. Aber die Existenz eines Cartesianischen Betrügergotts verneinen wir ja erst recht.

Descartes findet dann aber doch etwas, was vom skeptizistischen Zweifel ausgenommen ist (cf. seine 2. Meditation); ich kann mich nicht über *alles* täuschen. Denn bei allem Nachdenken über Betrügergötter, Sinnestäuschungen und anderem Rüstzeug aus dem skeptizistischen Arsenal muß ich doch stets feststellen, daß ich jetzt denke. Ich denke jetzt z.B., daß ein Betrügergott meinen Glauben, daß 2+2=4, unterminieren könnte. Aber ich kann nicht anzweifeln – und das ist der entscheidende Punkt –, daß *ich jetzt denke.*

In Descartes' Händen ist dieses Cogito, dieses „ich denke", der Ausgangspunkt für ein ganzes System. Zunächst einmal folgert er aus dem Umstand, daß er, Descartes, denkt, daß er selbst auch existiert – das berühmte „Ich denke, also bin ich". Dann setzt er – in der 3. Meditation – ein etwas merkwürdiges Kausalprinzip ein, wonach alles, was existiert, aus einer Ursache heraus existiert, wobei die Ursache in einem metaphysisch eigenartigen und explikationsbedürftigen Sinn immer 'größer' sein muß als ihre Wirkung. Das ermöglicht ihm, eine kausale Variante des ontologischen Got-

---

[5] Diese Geschichte stammt aus Putnam (1981).

[6] Auch Putnam selbst ist kein Skeptiker; der Kern seiner Diskussion der Hirne im Topf besteht in dem Versuch nachzuweisen, daß wir gerade keine solchen Hirne im Tank sein können. Cf. dazu weiter unten S.22f., wo sein Argument kurz dargestellt wird.

tesbeweises zu führen. Der so bewiesene Gott aber ist kein Betrüger, sondern allgütig; als ein solcher sorgt er dafür, daß in der Welt alles seine Ordnung hat und sich die armen Erkenntnissubjekte nicht grundsätzlich darüber täuschen, was um sie herum vorgeht. Es ist an dieser Stelle ganz besonders zu betonen, daß diese Skizze des Arguments aus den *Meditationen* nur eine Andeutung von Descartes' Argumentation vermitteln kann.[7]

Doch kehren wir lieber zu unserer eigenen Argumentationslinie zurück, die sich nur auf Descartes' ersten Argumentationsschritt verläßt. Drei Lehren lassen sich daraus ziehen:

(1) Es gibt offenbar unbezweifelbare Wahrheiten oder Gewißheiten; das Cogito ist so ein Fall.

(2) Es gibt so etwas wie synthetische Wahrheiten a priori.

(3) Der Skeptiker unterstellt, wir wüßten nur dann etwas, wenn wir uns dessen absolut sicher sind. Aber nicht jedes Wissen muß a priori sein. Wissen darf nicht mit Gewißheit identifiziert werden.

Das bedarf einiger Erläuterungen. (1) ist klarerweise das Ergebnis von Descartes' Reaktion auf den skeptischen Einwand: des Cogitos sind wir uns über jeden Zweifel hinaus sicher. (3) läßt sich nur mit (2) verstehen; und dazu sind erst einmal Begriffe zu klären. Was heißt „synthetisch", was heißt „a priori"?

Den Begriff einer synthetischen Wahrheit erklärt man am besten, indem man den zu „synthetisch" konträren Begriff „analytisch" erklärt. Der Einfachheit halber setzen wir „analytisch" gleich mit „analytisch wahr". Natürlich gibt es auch analytische Falschheiten, aber deren Explikation sollte nach den folgenden Ausführungen klar sein. Also: Eine Aussage ist genau dann analytisch, wenn sie allein aufgrund der Bedeutung der in ihr vorkommenden Ausdrücke und Wörter wahr ist; so lautet die Standarddefinition. Logische Wahrheiten und Begriffsdefinitionen sind die klassischen Beispiele für analytische Wahrheiten, z.B. „es regnet oder es regnet nicht", „Junggesellen sind unverheiratete Männer". Synthetisch sind nun alle die Aussagen, die weder analytisch wahr noch analytisch falsch sind. Das bedeutet, daß alle empirischen

---

[7] Angesichts dieser Karikatur sind ein paar Hinweise auf Sekundärliteratur zu Descartes nur angemessen. Gute Darstellungen findet man bei Röd (1982) und Williams (1978). Eine ältere, aber wichtige Aufsatzsammlung ist Doney (1968); darin sei besonders auf den Aufsatz von Hintikka verwiesen, der die spezifische Rolle des Cogito-Arguments untersucht.

Aussagen synthetisch sind, z.B. „es regnet“. Diese Terminologie hat bei Kant erstmals große Bedeutung gewonnen; seitdem ist ein genaues Verständnis des Analytizitätsbegriffs freilich eines der schwierigsten philosophischen Themen.

Kant hat zudem die Rede von „a priori“ und seinem Gegenbegriff „a posteriori“ eingeführt. Überzeugungen a priori liegen begrifflich vor aller Erfahrung. Sie schränken den Bereich möglicher Erfahrung nicht ein. Für sie ist wesentlich, daß sie *sich nicht als falsch herausstellen können*. Man beachte, daß das eine andere Behauptung ist als: *sie können nicht falsch sein*. Einfache Beispiele für Aussagen oder Überzeugungen, die sich nicht als falsch herausstellen können und damit a priori sind, lauten: „Ich bin jetzt hier“, „Ich existiere jetzt“.[8]

Die Begriffe der Analytizität und der Apriorizität muß man freilich auseinanderhalten. Natürlich sind analytische Wahrheiten auch a priori. Die logischen Positivisten meinten zu Beginn dieses Jahrhunderts, daß die umgekehrte Richtung ebenso gelte: Wahrheiten a priori sind auch analytisch. Doch das scheint falsch zu sein. Man sehe: „Ich existiere jetzt“ könnte falsch sein, kann sich aber für mich nicht als falsch herausstellen; es ist also nicht analytisch wahr, daß ich jetzt existiere. Damit haben wir ein Gegenbeispiel gegen die umgekehrte Richtung und zugleich ein Beispiel für eine synthetische Aussage a priori; ein und dieselbe Aussage demonstriert die obigen Behauptungen (1) und (2). Diese Aussage ist vielleicht nicht besonders aufregend; bei Kant gibt es viel spannendere Aussagen, denen er den Status synthetischer Wahrheiten a priori zuspricht. Wie es aber oft so geht, sind die interessanteren Fälle auch viel umstrittener.[9]

Wahrheiten a priori sind also unbezweifelbar (rational unbezweifelbar, heißt das stets). Ich denke, daß auch die Umkehrung gilt, daß also die Wahrheiten a priori als einzige rational unbezweifelbar sind. Aber das erforderte längere Argumentationen angesichts solcher Beispiele wie „mein Rücken schmerzt gerade“, welche doch auch

---

[8] Die meines Erachtens beste theoretische Beschreibung dieser Unterscheidungen in einem sprachphilosophischen Kontext findet man in Kaplan (1977), in Almog, Perry & Wettstein (1989). Vgl. auch Haas-Spohn (1995).

[9] In Kant (1781) finden sich im Abschnitt V der Einleitung der zweiten Auflage (B14ff.) verschiedene Beispiele dafür, was Kant alles als synthetische Urteile a priori ansieht: etwa die Aussagen der reinen Mathematik (die die Geometrie einschließt), oder bestimmte physikalische Gesetze, wie Newtons Prinzip „actio gleich reactio“. In der positivistischen Tradition des 20. Jahrhunderts. bestand dagegen die Tendenz, mathematische Aussagen auf logische zurückzuführen, die dann als analytisch anzusehen wären.

unbezweifelbar scheinen, aber sicherlich nicht a priori sind. In den Kapiteln 5 und 6 werden die Gründe für diese Umkehrung etwas klarer.

Wie dem auch sei, nunmehr läßt sich auch die Behauptung (3) einsichtig machen. Denn der Cartesische Skeptiker untergräbt ja unsere Wissensansprüche dadurch, daß er dauernd fragt, ob wir uns dessen, was wir wüßten, wirklich ganz sicher seien. Unsere Antwort kann dann immer nur sein: nein. Und so scheinen wir am Ende nichts (oder fast nichts) mehr zu wissen. Doch wenn das die Konsequenz wäre, so muß die Schlußfolgerung umgekehrt sein, daß die Voraussetzung der skeptizistischen Frage, Wissen sei maximal gewiß, einfach falsch ist. Daß dies eine Überforderung des Wissensbegriffs wäre, sieht man besonders deutlich, wenn man meine obige Behauptung akzeptiert, daß Wahrheiten a priori als einzige maximal gewiß sind. Denn dann implizierte die Voraussetzung des Skeptikers, daß man nur Wahrheiten a priori wissen kann – was offensichtlich absurd ist; die meisten Aussagen, um deren Wahrheit wir wissen, sind synthetische Aussagen a posteriori. Gibt man diese Voraussetzung auf, so bringt uns der Cartesische Skeptiker zunächst einmal nicht mehr in Verlegenheit.

Wir haben oben zwei mögliche Sorten skeptizistischer Einwände unterschieden, je nachdem, ob sie sich gegen den Wahrheitsanspruch oder gegen den Begründetheitsanspruch richten. David Hume soll nun als der historische Protagonist für diese zweite Art von Skepsis dienen, die die guten Gründe für unsere Wissensansprüche in Zweifel zieht. Seine Skepsis richtet sich gegen unseren Kausalitätsbegriff und gegen induktive Verfahren zum Erkenntnisgewinn. Wie noch zu sehen sein wird, hängen diese beiden Aspekte in Humes Werk eng miteinander zusammen. Auch Hume neigt dazu, Wissen mit Gewißheit zu identifizieren. Wissen im eigentlichen Sinn kann sich für ihn nur auf demonstrierbare Gewißheiten beziehen: auf mathematische Wahrheiten, die beweisbar sind oder introspektive Gewißheiten, also etwa Schmerzempfindungen. Alles andere ist für Hume kein Wissen in diesem engeren Sinn, sondern liegt im Bereich des Wahrscheinlichen.[10] Die skeptizistischen Schwierigkeiten, die er mit diesem Bereich hat, beziehen sich aber nicht einfach darauf, daß es dort im engeren Sinne nichts zu wissen gäbe, sondern gehen weiter:

Hume beginnt mit einem Bild vom Aufbau der empirischen Erkenntnis, das man durchaus als naiv bezeichnen darf. Danach besteht die Basis der Erkenntnis aus unmittelbar gegebenen Erfahrungen. Aus diesen grundlegenden Erfahrungen ziehen wir

[10] Humes klassische Schriften sind (1739) und (1748). Als Sekundärliteratur sei verwiesen auf Craig (1979), Bennett (1971) und Stroud (1977).

dann mehr oder weniger zwingende Rückschlüsse über die Beschaffenheit der Welt. Hume illustriert das an einem Beispiel. Er sitzt in seinem Arbeitszimmer und hört die Treppe knarren – das ist die unmittelbare Erfahrung. Weil zu dieser Stunde immer der Postbote zu kommen pflegt, bildet er auf der Basis dieser Erfahrung (und natürlich einer Menge weiterer Zusatzannahmen) die Erwartung, daß jetzt gleich der Postbote zu ihm kommt und ihm einen Brief von einem Freund bringt, welchen er, Hume, schon lang erwartet (cf. Humes 1739, I.IV.II). Man kann das Bild auch mit moderneren Beispielen illustrieren. Da sitzen die Hochenergiephysiker am DESY in Hamburg und suchen nach irgendwelchen Elementarteilchen. Sie erhalten Bilder aus den Detektoren, auf denen irgendwelche Streifen zu sehen sind. Aus diesem Datenmaterial folgern sie dann unter Zuhilfenahme ihrer Theorien, daß da infolge der herbeigeführten Kollision diese und jene Teilchen entstanden sind, die infolge physikalischer Gesetzmäßigkeiten diese und jene Spuren hinterlassen haben.

Aus diesem naiven Bild ergeben sich mithin zwei Probleme: das Basisproblem und das Induktionsproblem. Das Basisproblem besteht in der Frage, was denn die unmittelbar gegebenen Erfahrungen sind. Zwei Antworten wurden im Laufe der Geschichte immer wieder gegeben. Nach der einen besteht die Basis in beobachtbaren Gegenständen, die beobachtbare Eigenschaften aufweisen. Das liefert eine physikalistische Basis, die den Vorteil hat, daß sie intersubjektiv verfügbar ist. Allerdings sind solche Basiseinheiten nicht unmittelbar gewiß. Das ist der Vorteil der üblicherweise vertretenen Alternativposition: hier sind die Basiseinheiten die eigenen Sinnesdaten, also etwa der Grün-Eindruck, den man beim Blick auf die Tafel hat. Die sind unmittelbar gewiß, aber eben nicht mehr so ohne weiteres intersubjektiv verfügbar. Das Induktionsproblem besteht in der Frage: Wie kommen wir von der Basis zu all den möglichen Erweiterungen der Erkenntnis, und wie lassen sich diese Erweiterungen rechtfertigen? Hume ist Skeptiker bezüglich beider Probleme; ich befasse mich hier aber hauptsächlich mit dem Induktionsproblem.

Charakteristisch für Humes Philosophie ist die enge Verknüpfung seiner Skepsis hinsichtlich der Induktion mit seiner Analyse der Kausalität. Denn für ihn beruht alles induktive Schließen auf kausalen Schlüssen. Vom Auftreten einer Ursache schließen wir in der Regel auf Wirkungen dieser Ursache und umgekehrt von einer bestimmten Wirkung zurück auf die Ereignisse, die sie verursachten. Wenn Hume also die Treppe knarren hört, schließt er aus dieser Erfahrung auf ihre Ursache: daß der Postbote kommt – zum Beispiel, weil um diese Zeit immer der Postbote kommt. Und wenn er sieht, daß das Feuer im Ofen entfacht ist, schließt er darauf, daß es jetzt im Zimmer wärmer wird. Indem man diese Schlüsse immer weiter hintereinander ausführt, gewinnt man schließlich, ausgehend von der Basis, sein gesamtes empirisches Weltbild. (Wie man auf diesem Wege zu allgemeinen Gesetzen gelangt, ist nicht klar; aber zumindest was unsere Überzeugungen über einzelne Ereignisse anlangt, scheint Humes Position durchaus vertretbar.)

Unter diesen Voraussetzungen verlagert sich das Problem der Rechtfertigung induktiven Schließens auf die Rechtfertigung kausaler Schlüsse. Die traditionelle und bis heute immer wieder vertretene Idee besteht darin, daß zwischen Ursachen und Wirkungen eine Art notwendiger Beziehung besteht. Diese Idee ist Zielscheibe von Humes Kritik. Seine Analyse der Kausalrelation besagt zunächst:

> Ursache und Wirkung müssen tatsächlich vorliegen;
> die Ursache geht der Wirkung zeitlich voraus;
> Ursache und Wirkung müssen räumlich benachbart sein.[11]

Zu diesen allgemein anerkannten Bedingungen tritt üblicherweise noch als vierte hinzu, daß es eine notwendige Beziehung zwischen Ursache und Wirkung gibt. Aber dagegen erhebt Hume Protest. Wenn man sich einen typischen Fall einer Ursache-Wirkungs-Beziehung ansieht, also etwa die stets und immer wieder benutzte Geschichte von der Billardkugel, der auf einen anderen trifft und diesen in Bewegung versetzt, dann ist da keine notwendige Beziehung wahrzunehmen. Sie ist nicht zu sehen. Zudem ist z.B. die Vorstellung, daß die eine Kugel auf die andere trifft und beide dann still liegenbleiben, durchaus konsistent. De facto ist die Welt nicht so beschaffen; gemäß physikalischer Gesetze, insbesondere dem Satz von der Impulserhaltung, wird sich natürlich die zweite Kugel sich fortbewegen, nachdem der erste sie getroffen hat. Aber das ist keine *logische* Notwendigkeit. Und es sieht ganz so aus, als ob Hume, wenn er von Notwendigkeit spricht, immer logische Notwendigkeit im Auge hat. Ein zwar mit Vorsicht zu betrachtender, aber im allgemeinen doch zuverlässiger Test für die logische Möglichkeit eines Sachverhaltes besteht darin, sich zu fragen, ob man sich das Bestehen dieses Sachverhaltes konsistent vorstellen kann. Und das geht im Fall der Billardkugeln, die kollidieren und sofort liegenbleiben. Damit aber zeigt sich, daß kausale Relationen nicht mit logischer Notwendigkeit bestehen.

---

[11] „Ursache" und „Wirkung" sind hier als Bezeichnungen für einzelne Ereignisse zu verstehen, also für jeweils konkrete Vorkommnisse von etwas, was man als Ursache oder als Wirkung bezeichnen kann. In der dritten Bedingung geht es, genauer gesagt, nur um direkte Kausalverhältnisse.

Nimmt man die Geltung der tatsächlichen Naturgesetze an, dann folgt natürlich, daß die zweite Kugel sich bewegt. Aber woher haben wir etwa den Impulserhaltungssatz? Er ist ein allgemeiner Satz, der selber wieder von irgendwelchen Basisannahmen aus induktiv gewonnen werden muß. Und wenn schon für seine Einzelfälle eine Rechtfertigung nicht zu sehen ist – der Verweis auf kausale Notwendigkeit hilft jedenfalls nicht weiter –, so erst recht nicht für die Verallgemeinerung. Aus der Perspektive unserer Tage liegt freilich ein Einwand schon auf der Hand: Hume hat offenbar einen zu engen Begriff von Notwendigkeit unterstellt, wenn er nur logische Notwendigkeit für zulässig erachtet. Allerdings bleibt damit die positive Aufgabe im Raume stehen, einen Begriff von Naturnotwendigkeit zu explizieren, der das leisten kann, was Hume für unmöglich hielt.

Humes eigene Lösung des Problems bestand darin, darauf zu verweisen, daß wir einfach die Erwartung haben, daß der Ursache die Wirkung folgt. In uns laufen hier automatisierte, aber natürliche Prozesse ab. Hinter dem Begriff einer kausalen Notwendigkeit verbirgt sich für Hume nichts weiter als eine Denkgewohnheit, eine nützliche zwar, aber keine, die mit einem handfest metaphysischen Anspruch einhergehen könnte. Aber so klingt das einfach nach dem Eingeständnis, daß es eben keine Rechtfertigung für unser induktives Schließen gibt.

Wie nun bei dem skeptizistischen Angriff auf den Wahrheitsanspruch unserer Überzeugungen eine Verwechslung von Wissen mit Gewißheit die Argumentation beförderte, so scheint hier der Haken die Voraussetzung zu sein, man könne Begründungen nur dann akzeptieren, wenn sie absolut zwingend sind. Aber dem ist nicht so. Aus der Basis erhält man mit logisch zwingenden, also deduktiven Schlüssen eben keine Erweiterung der Erkenntnis. Daher liegt es nahe, zum Zwecke der Vermeidung der Begründungsskepsis den Begründungsbegriff abzuschwächen.

In der Tat sollte es einen nicht wundern, daß man hier Probleme bekommt: Deduktive Schlüsse sind natürlich wahrheitserhaltend (was heißt, daß mit den Prämissen auch die Konklusion wahr ist), und sie sind die einzige Art von Schlüssen, für die man das zwingend demonstrieren kann. Aber im Falle des Induktionsproblems geht es gerade um Erweiterungsschlüsse, also um Folgerungsbeziehungen, in denen der intuitive Gehalt der Folgerungen über den der Prämissen hinausreicht. Daß ein deduktiver Begründungsbegriff für die Explikation eines solchen Projekts unangemessen ist, ist dann nicht weiter erstaunlich.[12]

---

[12] Stegmüller (1975) macht das sehr deutlich. Eine gute Diskussion der Probleme rund ums induk-

An dieser Stelle muß man gegenüber Hume allerdings vorsichtig sein. Denn es ist eigentlich nicht klar, was er als eine gute Begründung ansieht. Unsere bisherige Diskussion verstand ihn, wenn man so will, unfreundlich, indem sie ihm einen an der deduktiven Logik orientierten Begründungsbegriff unterstellt. Das muß man nicht tun; man kann Hume auch anders deuten und seine skeptische Lösung selber als eine Art von pragmatischer Begründungsstrategie verstehen. Wie schon für die obige Diskussion von Descartes gilt auch hier: unter exegetischen Vorzeichen ist alles hier mit Vorsicht zu betrachten.

Humes Skeptizismus richtet sich aber auch gegen die Basis. Hier tritt folgendes Problem auf. Nehmen wir an, die Basis sei sicher, sie bestehe aus elementaren Erfahrungen, deren wir uns gewiß sind. Dann muß man diese elementaren Erfahrungen vermutlich phänomenalistisch auffassen, da man allenfalls Sinneseindrücke als absolut gewiß betrachten kann. Doch wie schließen wir nun von diesen Sinneseindrücken auf eine objektive Realität außer uns? An der Existenz einer solchen objektiven Realität zweifeln wir ja im allgemeinen nicht, und wir nehmen an, daß die Außenwelt durch eine von uns unabhängige, kontinuierliche Existenz ausgezeichnet ist. Unter den vorher gemachten phänomenalistischen Vorgaben kann der Schluß auf ihre Existenz wiederum nur induktiv sein. Doch erklärt Hume nun dem Leser, daß das nicht der übliche Sinn von „induktiv" ist. Der betrifft das Erschließen von Ursache-Wirkungs-Beziehungen, um die es aber momentan nicht geht. Stattdessen kommt hier ein Prinzip der Kontinuität und Kohärenz der Gegenstände unserer Sinneseindrücke ins Spiel, auf dessen Basis dann auf die unabhängige Existenz der Gegenstände geschlossen wird. Humes Diskussion dieser Frage (1739, Teil I.IV.II) ist langwierig und kommt eigentlich zu keinem Resultat, welches über eine skeptische Position hinausführte. Für Kant ist es der „Skandal der Philosophie", daß ein zureichender Nachweis der Existenz der Außenwelt bisher nicht gelungen sei; und er versucht dem Abhilfe zu leisten. In der Tat kann man die theoretische Philosophie Kants zu einem großen Teil als einen Versuch zur Beantwortung der von Hume hier aufgeworfenen Probleme betrachten.[13]

Wie auch immer, die Skepsis hinsichtlich der Möglichkeit guter Gründe läßt sich noch verschärfen. Die eindrücklichste Variante kommt aus dem 20. Jahrhundert; zu

---

tive Schließen findet man auch bei Salmon (1967).

[13] An dieser Stelle ist vor allem die *Widerlegung des Idealismus* aus der zweiten Auflage von Kants *Kritik der reinen Vernunft* relevant; cf. Kant (1781) B 274-279. Zu Humes Diskussion der induktiven Schlüsse auf die unabhängige Existenz externer Gegenstände cf. auch die einschlägige Diskussion in Bennett (1971).

finden ist sie in dem Büchlein *Fact, Fiction and Forecast* des amerikanischen Philosophen Nelson Goodman.[14] Goodman betrachtet nicht wie Hume Überzeugungen über zukünftige oder sonstwie jedenfalls nicht direkt beobachtete Ereignisse, sondern unsere allgemeinen Überzeugungen. Wie gelangen wir zu ihnen? Da scheint es eine ganz einfache Antwort zu geben: so war's halt immer schon. Goodmans mittlerweile klassisches Beispiel besteht in der allgemeinen Hypothese:

> Alle Smaragde sind grün.

Diese akzeptieren wir aufgrund, wie es heißt, enumerativer Induktion, d.h. einfach weil alle Smaragde, die wir bisher gefunden haben, tatsächlich grün waren und keiner eine andere Farbe hatte. Warum soll sich daran etwas ändern?[15]

Goodman hat nun einen Skeptiker erfunden, der das ganz anders sieht. Er erklärt, daß die korrekte Verallgemeinerung lautet:

> Alle Smaragde sind grot,

wobei das Prädikat „ist grot" wie folgt definiert ist:

> x ist grot genau dann, wenn[16] x vor dem 1.1. 2000 erstmals untersucht wurde und grün ist oder x nach dem 1.1. 2000 erstmals untersucht wurde und rot ist.

Nach dieser Definition sind alle bisher untersuchten Smaragde in der Tat grot. Warum erwarten wir, daß die Smaragde in der weiteren Zukunft, also insbesondere auch im nächsten Jahrhundert, grün sein werden und nicht vielmehr grot? Beide Hypothesen schließen sich aus: auf der Basis der ersten, der normalen, erwarten wir, daß der erste im nächsten Jahrhundert untersuchte Smaragd grün sein wird, auf der Grundlage der anderen, daß er rot sein wird. (Am 1.1. selber untersucht keiner Smaragde.) Aber beide Hypothesen sind so weit gleich gut begründet. Und da „x ist gut begründet" auf jeden Fall heißen sollte, daß x besser begründet ist als konkurrierende Alternativen,

---

[14] Literaturhinweise hierzu: Goodman (1955). Goodman hat sich weiter zu dem Thema ausgelassen in (1972). Diskussionen seines Arguments findet man z.B. in Stegmüller (1970ff.) Skyrms (1975).

[15] Wie so oft bei philosophischen Beispielen schlägt einem die Realität dann doch wieder ein Schnippchen. Es gibt eine Art von Smaragden, die man „Alexandriten" nennt. Chemisch werden sie unter die Smaragde eingeordnet; bei hellem Tageslicht sehen sie auch grün aus, in der Dämmerung oder bei Kerzenlicht erscheinen sie aber rot. Im folgenden vergessen wir diese Alexandriten besser wieder.

[16] Wir werden „genau dann, wenn" in Zukunft als „gdw." abkürzen.

bedeutet das, daß unsere Ausgangshypothese, daß alle Smaragde grün sind, nicht gut begründet ist. In Goodmans Terminologie läßt sich das damit aufgetretene Problem so beschreiben: manche Prädikate sind projizierbar, d.h. sie eignen sich für die Schlußweise der enumerativen Induktion, und manche sind es nicht; z.B. ist „grün" projizierbar, „grot" nicht. Doch wie läßt sich diese Unterscheidung explizieren und wie rechtfertigen, ein Prädikat so oder so zu klassifizieren?

Welch eine verrückte Idee, wird nun mancher sagen. Es sollte doch schlicht unzulässig sein, so seltsame Prädikate zu verwenden, in deren Definition ein völlig willkürlicher Bezug auf bestimmte Zeitpunkte auftritt. Man sollte fordern, daß wissenschaftlich akzeptable, projizierbare Prädikate rein qualitativ sind. „Grün" ist rein qualitativ, „grot" nicht.

Doch überzeugt das Goodmans Skeptiker gar nicht. Es ist seiner Ansicht nach vielmehr „grün", welches das seltsame, nicht qualitative Prädikat ist. Zum Nachweis erläutert er uns sein Prädikat „ist rün":

> *x* ist rün gdw. *x* vor dem 1.1.2000 erstmals untersucht wurde und rot ist oder *x* nach dem 1.1.2000 erstmals untersucht wurde und grün ist.

Die Definitionen von „grot" und „rün", so sagt der Skeptiker, sehen nur für Leute nicht-qualitativ aus, die so seltsame Prädikate wie „grün" projizieren. Aber in Wahrheit, fährt er fort, ist „grün" nicht-qualitativ; „grün" ist für ihn umgekehrt in der folgenden Weise zu definieren:

> *x* ist grün gdw. *x* vor dem 1.1.2000 erstmals untersucht wurde und grot ist oder *x* nach dem 1.1.2000 erstmals untersucht wurde und rün ist.

Mit dieser Methode, einem das Wort im Munde umzudrehen, legt Goodmans Induktionsskeptiker den Gedanken nahe, daß die Frage, ob ein Prädikat qualitativ ist oder nicht, nur relativ zu einer gegebenen Sprache zu entscheiden ist. Allemal hilft infolge der aufgezeigten Symmetrie der Verweis auf die Zulässigkeit nur rein qualitativer Prädikate nicht weiter.[17]

Diese Überlegungen verschärfen die Humesche Induktionsskepsis. Auch Goodman selbst schlägt als Ausweg nur vor, auf unsere bisherige Praxis zu rekurrieren.

[17] Viel Scharfsinn wurde darauf verwandt, unsere normale Position als die bessere auszuzeichnen; vgl. etwa die genannte Literatur. Doch scheint mir, daß sich die Symmetrie der normalen und der absurden Position nicht aufbrechen läßt.

Projizierbar sind, so Goodman, die Prädikate, die bereits in unserer bisherigen Praxis gut verankert sind. Das bringt mit sich, daß der Begriff des guten Grundes sich nur in Bezug auf die gesamte Praxis des Begründens explizieren läßt, weil jede weitere Verteidigung eines Teils unserer Praxis sich auf alle möglichen anderen Teile unserer Praxis berufen muß.

Diese Beispiele zeigen, daß eine gute Lehre, die man aus den skeptizistischen Einwänden ziehen kann, darin besteht, die Ansprüche an die beiden erkenntnistheoretisch zentralen Begriffe, Wissen und Begründung, zurückzuschrauben. Bevor ich mich nun einer konkreteren Analyse dieser Begriffe in ihrer zurückgenommenen Form zuwende, gilt es noch ein paar Nachbetrachtungen zu verschiedenen anderen skeptizistischen Argumentationen anzustellen.[18]

Es gibt weitere Formen des Skeptizismus neben den bisher betrachteten. Die wichtigste ist der Bedeutungsskeptizismus. Bei den untersuchten Angriffen auf Wissens- oder Begründungsansprüche war immer unterstellt, daß der Inhalt der in Frage stehenden Überzeugungen klar ist. Aber auch das kann man natürlich in Zweifel ziehen. Descartes' Betrügergott ist eine Figur, die sich zu diesem Zweck einsetzen läßt. Böser Geist, der er ist, vermag er uns so zu verwirren, daß wir nicht einmal mehr wissen können, was wir und ob wir überhaupt etwas mit unseren Worten (gesprochenen wie gedachten) meinen.

Putnams Geschichte von den Hirnen im Topf exemplifiziert ebenfalls eine Form der Bedeutungsskepsis. Wie schon erwähnt, ist Putnams Beweisziel, daß wir keine Hirne im Topf sein können. Sehr vereinfacht läuft sein Argument darauf hinaus, daß aus dem Umstand, daß wir denken können, daß wir Hirne im Topf sind, sich ergibt, daß wir keine Hirne im Topf sind. Wären wir Hirne im Topf, könnten wir nicht denken, wir seien welche. Auch die Hirne im Topf (HITs) haben ihre Sprache. So mag ein HIT denken „Die Bäume sind grün“. Worauf bezieht sich „Baum“? In unserer Sprache bezieht sich „Baum“ auf *Bäume*, und dies darum, weil wir in kausalem Kontakt zu Bäumen stehen. Aber „Baum“ im Topf-Deutschen bezieht sich nicht auf Bäume, denn die armen HITs haben ja gar keinen Kontakt zu Bäumen, jedenfalls nicht in der Version von Putnams Geschichte, in der die Hirne immer schon in ihrem Topf saßen. Nun wissen wir aber, worauf sich unsere Begriffe beziehen, meint Put-

[18] Für eine keinesfalls umfassende, aber doch recht repräsentative Diskussion der skeptischen Thematik cf. Stroud (1984).

nam – wir wissen, daß unser „Baum" sich auf *Bäume* bezieht. „Baum" im Topf-Deutschen bezieht sich nicht auf Bäume, also sind wir keine HITs.

Das ist eine Schnellversion eines umstrittenen und seltsamen Arguments. Das Unbehagen, das man damit haben kann, läuft aber blitzartig auf ein Argument für den Bedeutungsskeptizismus hinaus. Wieso sollen wir wissen, daß unser Wort oder Begriff „Baum" sich auf Bäume bezieht, und nicht vielmehr auf Vorstellungsbilder von Bäumen, Cumputerimplementationen oder dergleichen? Nehmen wir an, wir wissen es nicht. Dann ist die Geschichte mit den HITs in ihrer Allgemeinheit geeignet, einen zum Skeptiker in Bezug auf die Bedeutung der eigenen Worte zu machen.[19]

Auch die Goodman-Beispiele kann man so weiterspinnen; das hat insbesondere Kripke in seinem Buch über Wittgenstein eindrücklich ausgeführt.[20] Woher wissen wir denn, wer wir wirklich sind? Sind wir die normalen, für die wir uns halten, oder sind wir etwa die seltsamen Grotianer? Das ist nicht zweifelsfrei klar. Angenommen, wir graben am 2.1.2000 den nächsten Smaragd aus, und er ist grün. Der Grotianer ist schockiert und will uns gerade zu unserem Sieg gratulieren. Da bemerkt er, daß wir auch ganz betrübt sind und ihm recht geben; unserem Verhalten anch zu schließen, haben wir fest damit gerechnet, daß der Smaragd eine andere Farbe hat, d.h. rot ist. Wäre es dann nicht angemessen zu sagen, daß unser Wort „grün" offenbar schon immer „grot" bedeutet hat? Und woher wissen wir, daß wir uns nicht so verhalten werden? So läßt sich Goodmans Induktionsskepsis in eine Bedeutungsskepsis wenden.

Es gibt, neben der von uns betrachteten Strategie der Abschwächung zentraler Begriffe, noch andere Ausweichstrategien gegen skeptische Einwände. Wie gesehen, zeichnet sich der Skeptiker dadurch aus, daß er mit dem Fragen nicht aufhört. Er könnte etwa auf die Argumentation mit den HITs so reagieren, daß er sagt, daß unsere Annahme, keine HITs zu sein, ein unaufdeckbarer Irrtum sei. Aber das ist doch merkwürdig – was soll es heißen, daß etwas ein prinzipiell unaufdeckbarer Irrtum ist? Manche Philosophen haben auf solche Ansichten mit dem Hinweis reagiert, daß man eine Behauptung nicht unabhängig von Methoden zur Überprüfung ihrer Wahrheit oder Falschheit verstehen könne. Diese Reaktion ist unter der Bezeichnung *Verifik-*

---

[19] Das ist auch von so manchem in der Diskussion von Putnams Argument so gesehen worden. Für einen besonders klaren Fall sehe man Brueckner (1986).

[20] Cf. Kripke (1982), der in Wittgensteins *Philosophischen Untersuchungen* einen umfassenden Bedeutungsskeptizismus am Werke sieht, zu dem er eben auch Goodman fortentwickelt. Kripkes Diskussion ist höchst anregend, wie immer man sie als Wittgenstein-Interpretation wertet.

ationstheorie der Bedeutung bekannt geworden. Akzeptiert man eine derartige Bedeutungstheorie, kann man dem Skeptiker entgegenhalten, daß er uns jegliche Methode zur Verifikation von Sätzen raubt und damit anfängt, nur noch unverständlich zu reden. Allerdings ist die Verifikationstheorie der Bedeutung auch umstritten und hat seit einiger Zeit keine gute Presse. Carl Gustav Hempel, der die Verifikationstheorie einst selbest vertreten hat, dokumentiert in einem schon klassisch gewordenen Aufsatz sehr schön die Vorteile und vor allem die Nachteile dieser Theorie.[21]

Eine andere Reaktion besteht in einer Unterscheidung von lokaler und globaler Skepsis. Der Skeptiker stellt fest, daß er jede einzelne Behauptung in Zweifel ziehen kann, und folgert daraus, daß er auch alle miteinander in Zweifel ziehen könne. Aber es ist alles andere als klar, daß das ein legitimer Schritt ist. So beinhaltet die Institution des Versprechens, wie wir sie alle kennen, daß es lokale Brüche von Versprechen geben kann. Aber wenn überhaupt kein Versprechen mehr gehalten wird, ist die Institution eo ipso verschwunden. Analog könnte man sagen, daß ein globaler Zweifel nicht einmal mehr ein Zweifel ist.[22] Allerdings hakt die Analogie ein wenig, denn Wahrheit ist nicht konventional, im Gegensatz zur Institution des Versprechens.

Schließlich ist noch die Reaktion von Donald Davidson erwähnenswert. Er beansprucht, ein prinzipielles Argument dafür zu geben, daß die Mehrzahl unserer Überzeugungen wahr sein muß. Das Argument erfolgt im Rahmen seiner Theorie der sogenannten radikalen Interpretation und läuft ganz kurz gesagt so: Wir sehen uns einem Individuum gegenüber, das in seiner natürlichen und sozialen Umwelt bestimmte Zeichen, Gesten und Geräusche produziert. Wir wollen dieses Wesen verstehen und seine Äußerungen interpretieren. (Diese Interpretation heißt eine radikale, weil sie nicht voraussetzt, daß wir mit diesem Individuum schon in einem gemeinsamen kulturellen Verständigungskontext stehen.) Nach Davidsons berühmten ‚Principle of Charity' kann nun eine Interpretation nur dann etwas taugen, wenn sie die meisten Überzeugungen des Interpretierten wahr macht; anders sind wir nicht in der Lage, ihn überhaupt zu verstehen. Das heißt aber, daß wir mit der Attitüde des Skeptikers ihn nicht verstehen und ihm nicht gerecht werden können. Damit erledigt sich der Bedeutungs- und der Wissensskeptizismus auf einen Schlag; die Bedeutung unserer Wörter sind so, daß die meisten unserer Überzeugungen wahr sind. Das läßt

[21] Cf. Hempel (1959).
[22] Cf. Wittgenstein (1984).

noch Raum für beliebigen lokalen Zweifel; aber der globale Zweifel wäre damit widerlegt.[23]

Bei diesem kursorischen Überblick über Argumentationen zur Skeptizismus-Problematik will ich es hier bewenden lassen, trotz ihrer schwer widerstehlichen Anziehungskraft. Aber es dürfte auch zu spüren sein, daß man wesentlich genauer in all diese Argumentationen hineingehen müßte, wenn man zu begründeten Bewertungen über sie kommen wollte. Zwei Punkte waren aber so schon deutlich; der Wissens- und der Begründungsskeptizismus zwingen uns dazu, diese beiden Begriffe genauer zu betrachten. Um den Wissensbegriff geht es im nächsten Kapitel; der Begründungsbegriff bedarf hingegen einiger Vorarbeiten und wird so erst im Kapitel 6 ausführlich thematisiert.

[23] Für den letzten Stand seiner Position s. Davidson (1990); ansonsten Davidson (1984).

3. Analysen des Wissensbegriffs

Wo der Skeptiker Wissensansprüche problematisiert, ist (zumindest unter analytischen Philosophen) die natürliche erste Frage, was es denn überhaupt heißt, daß jemand etwas weiß; der Begriff des Wissens ist erst genauer zu analysieren. Rudolf Carnap hat einst eine sehr nützliche Liste von Forderungen zusammengestellt, denen eine Begriffsanalyse oder Begriffsexplikation genügen sollte.[24] Sie lautet:

1. Das Explikat sollte dem Explikandum ähnlich sein.[25]
2. Das Explikat sollte klar und präzise sein.
3. Das Explikat sollte fruchtbar sein.
4. Das Explikat sollte, falls möglich, einfach sein.

Diese Kriterien sollen nun auch auf die Analyse des Wissensbegriffs angewandt werden.

Doch noch vor aller Analyse ist eine wichtige Unterscheidung zu beachten: die zwischen propositionalem Wissen und, wie man es nennen könnte, praktischem Wissen. Im Englischen läßt sie sich sehr einfach und deutlich ausdrücken als die Unterscheidung von „know how" und „know that". Unter den ersten Begriff fallen alle möglichen Fertigkeiten und Fähigkeiten. In der erkenntnistheoretischen Tradition richtete sich die Aufmerksamkeit jedoch fast ausschließlich auf den Begriff des propositionalen Wissens. Ein Philosoph wie Gilbert Ryle hat heftig gegen diese Tendenz protestiert.[26] Doch werde auch ich mich hier aufs „knowing that" beschränken.

Die Tradition der Analyse des Wissensbegriffs ist ehrfurchtgebietend alt. Platons *Theaitetos* kann man als den ersten Klassiker bezeichnen. Dort findet man nicht unbedingt eine Analyse in dem Stil, wie wir es mittlerweile gewöhnt sind; die meiste Zeit befaßt sich Platon mit bestimmten ontologischen Fragen und damit, wie es möglich ist, falsche Meinungen zu haben. Aber Sokrates stellt in diesem Dialog auch drei

[24] S. Carnap (1950).

[25] Das soll bedeuten, daß eine schlichte Nominaldefinition nicht ausreicht.

[26] Cf. Ryle (1949). Ryle ist eine der herausragenden Gestalten der sogenannten Philosophie der normalen Sprache. Für eine umfassende Darstellung dieser Philosophie s. von Savigny (1993).

Theorien zum Wissensbegriff vor, von denen eine für uns von Bedeutung ist. Wissen besteht danach in wahren Überzeugungen. So lautet eine erste Analyse wie folgt:

(a) x weiß, daß p, gdw. (1) p der Fall ist und (2) x glaubt, daß p.

Eine Begriffsanalyse soll notwendige und hinreichende Bedingungen für die Anwendbarkeit des zu analysierenden Begriffs liefern. Kein Zweifel, Platons Analyse liefert notwendige Bedingungen. Man kann überhaupt nur wahre Dinge wissen. Und wenn man etwas weiß, dann glaubt man es auch – in dem halb-technischen Sinn dieser Begriffe, den sie in der Philosophie nun einmal haben. Es gibt gewiß umgangssprachliche Verwendungen von „wissen" und „glauben", die sie zu Gegensätzen machen; so aber werden sie nicht in der Philosophie verwendet.[27] Ferner gibt es in manchen Bereichen, etwa in der Wissenssoziologie oder in der KI-Forschung, eine Neigung, den Wissensbegriff so zu verwenden, daß Wahrheit nicht darin enthalten sein muß. Mir scheint damit aber die wesentliche Differenzierung zwischen Wissen und Glauben verlorenzugehen, weshalb ich diesen Sprachgebrauch ausdrücklich zurückweise.

Doch schon Platon hatte Vorbehalte gegen die Ansicht, man hätte mit der erwähnten Analyse bereits hinreichende Bedingungen. Er läßt Sokrates die Geschichte von einem Advokaten erzählen, der vor Gericht die Richter mit einer gut vorgetragenen Rede von der Wahrheit einer Aussage überzeugt, die zufälligerweise auch wahr ist. Aber es ist zweifelhaft, ob die Richter dann über Wissen verfügen (cf. *Theaitetos*, 201b/c). Noch klarere Fälle treten auf, wenn man Antworten errät. Man fragt mich, wie viele Fischotter in Deutschland in freier Wildbahn leben. Ich sage, daß es 350 sind. Zufällig stimmt die Zahl (hoffentlich sind es überhaupt noch so viele!). Habe ich damit gewußt, wie viele Fischotter es in Deutschland in freier Wildbahn noch gibt? Es scheint höchst unplausibel, dies zu bejahen.

Platon legt Sokrates eine andere Analyse in den Mund, die da modernisiert lautet:

(b) x weiß, daß p, gdw. x gesehen oder wahrgenommen hat, daß p.

---

[27] Es ist oft so, daß ein schwächerer Begriff als einen stärkeren Begriff ausschließend verstanden wird; z.B. ist man geneigt, aus der Aussage „einige Studenten sind fleißig" die Aussage „nicht alle Studenten sind fleißig" zu folgern. Das ist jedoch ein pragmatisches Phänomen. Wenn ich die stärkere Aussage für wahr hielte, so wäre es seltsam, wenn ich nur die schwächere Aussage machte. Wo ich nur die schwächere mache und mich in der Sache, über die ich rede, auskenne, darf man also unterstellen, daß ich die stärkere Aussage für falsch halte; das ist die Begründung der Folgerung. Dieser pragmatische Umstand bedeutet aber nicht, daß es in der wörtlichen Bedeutung von „einige" liege, daß es „nicht alle" impliziert. Dasselbe gilt mutatis mutandis für „glauben" und „wissen".

In der Tat führen unsere Wahrnehmungen typischerweise zu Wissen. Doch ist (b) eine Verschärfung von (a), die zu weit geht. Wir wissen einiges, was wir nicht gesehen oder wahrgenommen haben, jedenfalls dem üblichen Verständnis von „sehen“ oder „wahrnehmen“ zufolge. In welchem Sinn beruht mein Wissen, daß Elektronen geringere Masse haben als Neutronen, auf Wahrnehmung? Man könnte sagen, ich hätte es gelesen und damit natürlich auch wahrgenommen. Aber das ist nicht der übliche Sinn von „wahrnehmen“. Im besten Fall liefert (b) also eine hinreichende Bedingung.

So setzte sich schließlich die folgende Analyse durch, die Platon ebenfalls ganz am Ende des *Theaitetos* in anderer Ausdrucksweise nahelegte und die über 2000 Jahre den Standard vorgab:

(c) x weiß, daß p, gdw. (1) p ist der Fall,
(2) x glaubt, daß p,
(3) x ist in seiner Überzeugung, daß p, gerechtfertigt.

Von dieser Analyse mag man hoffen, daß sie wirklich notwendige und hinreichende Bedingungen liefert. Kaum einer zweifelt, daß sie notwendig sind. Aber ist sie auch hinreichend? Dieser Frage werde ich in diesem Kapitel hauptsächlich nachgehen.

Zuvor ist aber festzustellen, daß der Philosoph natürlich jede der drei Klauseln der momentan vorgeschlagenen Analyse noch weiter erläutern muß. Sofern er als Erkenntnistheoretiker spricht, braucht er zu (1) nichts zu sagen, wohl aber zu den beiden anderen Punkten. Für (2) muß er eine Theorie des Glaubensbegriffes erarbeiten; dazu folgt in den späteren Kapiteln mehr. Und für (3) muß er etwas über Begründungen und Rechtfertigungen sagen (wozu ihn ja auch die Skeptizismusproblematik schon aufgefordert hat). Darauf werde ich im Kapitel 6 näher eingehen. Doch scheinen mir einige grundlegenden Bemerkungen schon jetzt am Platze zu sein.

Die Diskussion von Humes Skeptizismus hatte ja deutlich gemacht, daß es nicht immer zwingende Begründungen geben muß. In unseren alltäglichen Redeweisen zeigt sich der Unterschied zwischen verschieden starken Weisen der Begründung und Rechtfertigung ganz deutlich. Es gibt Ausdrücke, die auf deduktive Zusammenhänge verweisen: Wir *folgern* einen Satz, *widerlegen* eine These, weisen die *Konsistenz* einer Theorie nach. Dem gegenüber steht ein Vokabular, das schwächere Beziehungen zum Ausdruck bringt: Ein Umstand *spricht für* eine Ansicht, eine Beobachtung *entkräftet* eine Hypothese, eine Überlegung *macht* eine Auffassung *plausibel* usw.

Wenn wir nun im Rahmen einer Analyse des Wissensbegriffs nach einer Theorie der Begründung Ausschau halten, dann geht es um diese schwächeren Begriffe. Der Grundbegriff einer derartigen Theorie ist offenbar relational:

(R) *A* ist ein Grund für *B*.

Diese Relation zwischen Sätzen, Sachverhalten, Überzeugungsinhalten oder ähnlichem gilt es im Rahmen einer solchen Theorie zu klären. Nun habe ich gerade in (R) einen relationalen Begriff als Grundbegriff angegeben; in Bedingung (3) der Wissensanalyse war hingegen der Begriff der Rechtfertigung nicht relational. Der Zusammenhang zwischen beidem liegt auf der Hand: Eine Überzeugung ist offenbar dann gerechtfertigt, wenn das Subjekt gute Gründe für sie hat; und gute Gründe sind solche, die ihrerseits gut begründet sind.

Mit dieser Erläuterung tut sich jedoch ein Problem auf: das sogenannte Begründungstrilemma. Gute Gründe sind solche, sagten wir eben, die ihrerseits gut begründet sind. Jetzt droht aber (a) ein Regreß oder (b) ein Zirkel oder (c) Dogmatismus. Denn entweder verweist die Rechtfertigung der Güte bestimmter Gründe auf immer andere Gründe, ohne je zu einem Ende zu kommen; das wäre ein unendlicher Regreß, durch den letztendlich nichts zu rechtfertigen ist. Oder man wird im Verlauf der Begründungskette irgendwann wieder einen Grund anführen, der man schon vorher angeführt hat; das wäre ein Zirkel, der anscheinend auch den Rechtfertigungsanspruch untergräbt. Oder man gelangt schließlich zu Behauptungen, die keiner weiteren Begründung mehr fähig sind und womöglich auch keine mehr nötig haben; diese scheinen also dogmatisch vom Himmel zu fallen.

Die drei Alternativen des Begründungstrilemmas sind erschöpfend. Zwar hat niemand den Regreß als akzeptabel ins Auge gefaßt; aber die beiden anderen Positionen spielen in der erkenntnistheoretischen Debatte unserer Tage eine große Rolle. Diejenigen, die die eben als Dogmatismus bezeichnete Reaktion für angemessen oder unausweichlich halten, bezeichnet man als *Fundamentalisten*.[28] Nach ihrer Auffassung gibt es in Form von Wahrnehmungen oder vielleicht auch irgendetwas anderem in der Tat eine Basis, deren Elemente sicher, evident oder jedenfalls nicht mehr begründungspflichtig sind. Das andere Lager umfaßt die *Kohärentisten*. Sie halten Zirkel, die in den Begründungsketten auftreten, gegebenenfalls für harmlos und unvermeid-

---

[28] Das englische Etikett „foundationalists“ ist weniger belastet, da man dort auch die „fundamentalists“ kennt; doch können wir im Deutschen den Unterschied nicht gut wiedergeben.

lich und lehnen die Idee ab, es gäbe eine feste Basis; auch Wahrnehmungen oder direkt der Wahrnehmung entnommene Überzeugungen unterliegen für sie einer gewissen Begründungspflicht. Doch liegt die Begründung einer jeden Überzeugung für sie darin, daß sie sich in ein kohärentes System von Überzeugungen einfügt. Damit übernehmen sie freilich die Aufgabe, den Begriff der Kohärenz genauer zu erläutern.[29]

Es kann hier nicht der Ort sein, eine umfassende Diskussion über Vor- und Nachteile der beiden Konzeptionen durchzuführen; vermutlich wird man beide miteinander verbinden müssen. Nicht verhehlen möchte ich jedoch, daß mir der Kohärentismus attraktiv erscheint, vor allem dann, wenn man in seinem Rahmen dem offenbar vorhandenen Sonderstatus von Wahrnehmungsüberzeugungen gerecht werden kann; wir werden darauf zurückkommen.

Diese kurzen Bemerkungen zeigen aber wieder eines deutlich: daß am Grunde der Wissenstheorie eine Theorie des Glaubensbegriffes stehen muß. Beide Konzeptionen, Fundamentalismus wie Kohärentismus, müssen etwas dazu sagen können, wie Überzeugungen im Lichte neuer Gründe verändert werden, warum manche Arten von Überzeugungen gegenüber Veränderungen größere Resistenz aufweisen sollen als andere usw. Und dazu sind wir wiederum auf eine Theorie des Glaubens sowohl in seinen statischen wie seinen dynamischen Aspekten verwiesen; ihre Zentralität tritt immer deutlicher hervor.

Doch wenden wir uns nun der Hauptfrage dieses Kapitels zu, nämlich ob die obige Analyse (c) des Wissensbegriffs bereits akzeptabel ist oder nicht. Die Diskussion, deren Verlauf ich im folgenden in Auszügen wiedergeben will, ist ein typisches Beispiel dafür, wie in der analytischen Philosophie oft vorgegangen wird. Man schlägt eine Analyse vor, sucht sie durch Gegenbeispiele zu erschüttern, verbessert die Analyse im Hinblick auf die anerkannten Gegenbeispiele, sucht erneut nach solchen usw. – bis man eine allgemein anerkannte Analyse gefunden hat, oder (was vermutlich öfter der Fall sein dürfte) bis die philosophische Gemeinschaft von der Debatte etwas ermüdet ist oder etwas Abstand zu ihr gewinnt und manche Detaildiskussion als in der einen oder anderen Weise irregeleitet erkennt – und dann ihren Scharfsinn anderen Analyseproblemen zuwendet. Debatten dieser Art gab es etwa im Anschluß an die Arbeiten von Hempel zum Begriff der wissenschaftlichen Erklärung oder in der

[29] Eine Monographie in der fundamentalistischen Tradition ist z.B. Moser (1989). Die kohärentistische Linie ist sehr gut durch BonJour (1985) vertreten; vgl. dazu auch Bender (1989).

Sprachphilosophie im Anschluß an Paul Grices Analyse des Begriffs des Meinens.[30] Mit dem Wissensbegriff sieht es nicht anders aus.[31]

Unsere zuletzt erreichte Analyse, deren Notwendigkeit anerkannt war, bei der aber noch zu überprüfen ist, ob sie hinreichend ist, lautete:

(c) *x* weiß, daß *p,* gdw. (1) *p* ist der Fall,
(2) *x* glaubt, daß *p*,
(3) *x* ist in seiner Überzeugung, daß *p*, gerechtfertigt.

Diese Analyse war allgemein anerkannt, bis im Jahr 1963 ein Philosoph mit einem drei Seiten langen Artikel, der wohl seine einzige Publikation bleiben sollte, sie erschütterte. Der Mann heißt Edmund Gettier; seine Leistung bestand darin, mittels Beispielen zu zeigen, daß die klassische Analyse nicht hinreichend ist.

Bevor wir Gettiers eigene Beispiele besprechen, soll aber anhand eines etwas einfacheren Beispiels von Roderick Chisholm die Art von Einwand präsentiert werden, die die ehrwürdige Theorie zu Fall bringt. Ich stehe auf einer Wiese und glaube, daß da auch ein Schaf auf der Wiese ist. Ich sehe es – jedenfalls meine ich das –, und da ich normalerweise Schafe gut zu erkennen vermag, habe ich infolge meines Eindrucks gute Gründe für meine Überzeugung, daß da ein Schaf auf der Wiese ist. Tatsächlich aber ist das Tier, das ich für ein Schaf halte, ein Bedlington-Terrier, diese seltsame Rasse, die man wirklich oft erst auf den zweiten oder dritten Blick als Hunde und nicht als Schafe identifiziert. Wie es der Zufall will, befindet sich weiter hinten auf der Wiese ein Schaf, eine dunkle Heidschnucke, die ich aber auch nicht richtig erkenne. Also: es ist ein Schaf auf der Wiese, ich glaube, daß ein Schaf auf der Wiese ist und ich habe gute Gründe dafür, zu glauben, daß ein Schaf auf der Wiese ist. Aber es ist doch intuitiv klar, daß ich nicht weiß, daß ein Schaf auf der Wiese ist.

Gettiers eigene Beispiele sind etwas verwickelter.[32] Da wartet Schmid auf das Ergebnis eines Vorstellungsgesprächs, zu dem er und Maier eingeladen wurden. Aus dem Verlauf des Gesprächs für ihn und aus der Plauderei mit Maier hat Schmid gute

---

[30] Die wichtigsten Arbeiten Hempels zum Thema „wissenschaftliche Erklärung“ finden sich in (1965). Einen im wesentlichen repräsentativen und umfassenden Überblick über die Debatte gibt Salmon (1989).– Die Arbeiten von Grice liegen zusammengefaßt vor in Grice (1989). Die Debatte dazu wird dokumentiert z.B. in Meggle (1979).

[31] Wichtige Aufsätze in deutscher Übersetzung enthält Bieri (1987) Weitere Literatur: Pappas und Swain (1978); Chisholm (1989), Goldman (1986) und Lehrer (1990).

[32] Cf. Gettier (1963).

Gründe für die Überzeugung gewonnen, daß Maier die Stelle kriegt und 10 Mark in der Hosentasche hat. Deshalb glaubt Schmid auch, daß Maier die Stelle kriegt *und* 10 Mark in der Hosentasche hat. Nun kann Schmid folgern, daß derjenige, welcher die Stelle kriegt, 10 Mark in der Hosentasche hat. Diese Folgerung aus Schmids Überzeugungen ist ebenfalls gut begründet. Schmid hat also die gut begründete Überzeugung, daß derjenige, der die Stelle kriegt, 10 Mark in der Hosentasche hat. Die Bedingungen (2) und (3) der Wissensanalyse (c) sind erfüllt. Und siehe da: (1) ist auch erfüllt; derjenige, der die Stelle kriegt, hat 10 Mark in der Hosentasche. Denn tatsächlich kriegt Schmid die Stelle, und – was er aber selbst nicht bemerkt hat – er hat auch 10 Mark in der Hosentasche. Unter diesen Umständen wäre es äußerst seltsam zu sagen, Schmid wüßte, daß derjenige, der die Stelle kriegt, 10 Mark in der Tasche hat.

Das andere Gegenbeispiel gegen die klassische Analyse läuft folgendermaßen: Schmid glaubt, daß Maier einen Ford besitzt. Er hat gute Gründe für diese Überzeugung; so hat er Maier oft mit dem Auto fahren sehen, sah den Fahrzeugschein, in dem Maier als Besitzer eingetragen ist. Aber Maier hat mittlerweile keinen Ford mehr, was Schmid allerdings nicht weiß. Schmid kann aber nun aus „Maier besitzt einen Ford“ folgern „Maier besitzt einen Ford, oder Müller ist in Bielefeld bzw. Müller ist in Herford bzw. Müller ist in Gütersloh bzw. Müller ist in Detmold“. Schmid glaubt auch, wenn er rational ist, diese Folgerungen aus seiner ersten Überzeugung. Und da es sich dabei um schlichte logische Folgerungen handelt, sind sie ebenso gerechtfertigt wie die Ausgangsüberzeugung.[33] Nehmen wir nun an, Müller ist in Bielefeld; irgendwo muß er ja sein. Dann ist „Maier besitzt einen Ford oder Müller ist in Bielefeld“ wahr, Schmid glaubt, daß Maier einen Ford besitzt oder Müller in Bielefeld ist, und unseren Annahmen zufolge ist diese Überzeugung von Schmid gut begründet. Aber Schmid weiß nicht, daß Maier einen Ford besitzt oder Müller in Bielefeld ist. Daß Maier einen Ford besitzt, kann er nicht wissen, weil Maier gar keinen Ford mehr hat, und über Müller weiß er eigentlich nichts.

In diesen Beispielen wird stets eine Minimalrationalität der Protagonisten unterstellt, was aber harmlos ist. Zudem benutzt Gettier explizit das Prinzip, daß q gut begründet ist, falls p gut begründet ist und q eine unmittelbare logische Folgerung aus p ist. Auch diese Annahme erscheint plausibel; zudem funktioniert das Gegenbeispiel von Chisholm ja auch ohne diese Annahmen. Also zeigen all diese Beispiele in der Tat, daß etwas bei der Analyse des Wissensbegriffs fehlt.

[33] Aus p folgt $p \vee q$, für beliebige q. Und ein rationaler Akteur glaubt zumindest die unmittelbaren Folgerungen aus seinen Überzeugungen.

In allen Gegenbeispielen fällt eines auf: die Gründe des Protagonisten für seine Überzeugung leiten sich gerade nicht von der Wahrheit des geglaubten Sachverhalts her. Daß ein Schaf auf der Wiese steht, hat mit den Gründen für meine Überzeugung nichts zu tun. Daß Schmid glaubt, daß derjenige, der die Stelle kriegt, 10 Mark in der Hosentasche hat, hat nichts damit zu tun, daß Schmid die Stelle kriegt, und daß Schmid glaubt, daß Maier einen Ford hat oder Müller in Bielefeld ist, hat nichts mit Müller zu tun. Das legt den Verbesserungsvorschlag nahe, den Bedingungen der Wissensanalyse die folgende vierte hinzuzufügen:

(4C) alle Gründe, die x für die Überzeugung, daß p hat, sind wahr.[34]

Aber das stößt schon wieder auf Gegenbeispiele. So glaube ich, daß jemand hier im Hörsaal einen Ferrari besitzt. Meine Gründe dafür sind, daß da drüben Herr Protz sitzt, der mich gelegentlich mit einem Ferrari mitnimmt, mit seinem Auto auch anderweitig gerne angibt usw. Was ich jedoch nicht weiß, ist, daß er ihn gar nicht besitzt; er hat den Wagen nur geliehen. Was ich ebensowenig weiß, ist, daß sich im Hörsaal auch Herr Krämer befindet, der tatsächlich einen Ferrari besitzt, aber daraus ein Geheimnis macht. So weit haben wir ein Beispiel des alten Typs; die Bedingung (4) schließt nun aus, daß wir sagen, ich wüßte, daß einer hier im Raum einen Ferrari besitzt. Aber da ist auch noch Herr Reich im Hörsaal. Der besitzt einen Ferrari; und ich weiß, daß er einen Ferrari besitzt. Also weiß ich, daß einer im Hörsaal einen Ferrari besitzt. Ich habe viele gute Gründe für diese Überzeugung. Aber nicht alle sind wahr – einer war ja meine Überzeugung, daß Protz einen Ferrari besitzt, was nicht stimmt. Und damit schließt die neue Bedingung (4) aus, daß ich weiß, daß einer hier im Hörsaal einen Ferrari besitzt. Das ist ersichtlich unplausibel; ich weiß doch, daß Reich einen Ferrari hat. Offensichtlich ist die Bedingung (4) zu stark; sie schließt Fälle aus, die wir intuitiv als Fälle von Wissen klassifizieren würden.

Dafür, daß eine bestimmte Überzeugung Wissen darstellt, reicht es also offenbar hin, daß nur einige der Gründe für sie auch wahre Gründe sind; das gilt es nun genauer zu erfassen. Es gibt da zwei Denkschulen, die sich *Externalisten* und *Internalisten* nennen und die unterschiedlich auf diese Sachlage reagieren. Externalisten richten ihr Augenmerk auf die äußeren Umstände der Entstehung von Wissen; einer ihrer herausragenden Repräsentanten ist Alvin Goldman mit seiner Kausaltheorie des Wissens. Internalisten suchen dagegen nach weiteren internen Bedingungen im Subjekt selbst, die wissensstiftende Begründungen auszeichnen; zur Illustration will ich Ideen von

[34] Das „C“ rührt daher, daß Clark (1963) als erster diesen Vorschlag gemacht hat.

Keith Lehrer präsentieren, der einen Begriff der nicht entkräftbaren hinlänglichen Begründung zu entwickeln versucht.[35]

Betrachten wir zunächst die Theorie von Alvin Goldman. Auch er sucht nach einer vierten Bedingung für die Analyse. Die Gettier-Beispiele beruhen, wie schon gesagt, alle darauf, daß kein Zusammenhang zwischen der Wahrheit von *p* und den Gründen der Protagonisten, *p* zu glauben, besteht. Nach Goldman ist der wissensstiftende Zusammenhang gerade ein kausaler, und so formuliert er die folgende vierte Bedingung:

(4G) Es besteht ein kausaler Zusammenhang zwischen *p* und *x*' Überzeugung, daß *p*.

In allen Fällen lautet eine entscheidende Frage: wie kam die Meinung des Wissenssubjekts zustande, hat sie mit dem Bestehen der gewußten Tatsache zu tun oder nicht? Wenn das Bestehen der Tatsache eine kausale Rolle bei der Bildung der Meinung von *x* spielt, kann *x* wissen, daß *p*. Die Idee, Kausalität ins Spiel zu bringen, ist nicht neu. So hat man Wahrnehmungswissen schon immer über die Ursachen der Wahrnehmung zu analysieren versucht: der Umstand, daß vor mir eine Vase steht, ist Ursache meiner Wahrnehmungsüberzeugung, daß vor mir eine Vase steht; es ist die Vase selbst, die bei mir ein Netzhautbild verursacht, welches den Input für weitere Erkenntnisprozesse in mir bildet. Ein anderer Bereich, in dem Wissensansprüche mit Hilfe kausaler Begrifflichkeit expliziert werden, ist der der Erinnerung. Man möchte ja Erinnerungen von scheinbaren Erinnerungen unterscheiden, etwa von solchen Fällen, in denen ich nur glaube, mich an ein Ereignis zu erinnern, in Wahrheit aber nur noch Erzählungen davon kenne. In diesen Fällen ist es nicht das erinnerte Ereignis selbst, das meine Überzeugung verursacht hat; und damit kann diese Überzeugung nicht als eine Erinnerung klassifiziert werden.[36]

Wie steht Goldman zu den oben dargelegten Gegenbeispielen zur Analyse (c)? Seine Theorie kommt mit dem Schaf-Beispiel zurecht. Ich weiß nicht, daß ein Schaf

---

[35] Für die folgenden Überlegungen cf. die Artikel von Goldman und von Lehrer und Paxson jr. in Bieri (1987) sowie die oben erwähnten Bücher von Goldman und Lehrer, die repräsentativ für die beiden Denkschulen sein dürften. Noch ein terminologischer Hinweis: Die Bezeichnung von Denkschulen als *Externalisten* und *Internalisten* ist in verschiedenen Bereichen üblich. So bezeichnet man in der Philosophie des Geistes die Gegner des Prinzips des methodologischen Solipsismus zuweilen auch als Externalisten; cf. die Diskussion im Kapitel 7. Das ist eine völlig andere Verwendung als die derzeit betrachtete, die allein einer bestimmten Position in der Erkenntnistheorie zukommt; die verschiedenen als externalistisch bezeichneten Positionen sind voneinander weitgehend unabhängig.

[36] Klassiker dieser Form von Analyse sind bzüglich der Wahrnehmung Grice (1961), und hinsichtlich Erinnerungen Max Deutscher und C.B. Martin (1966).

auf der Wiese ist, weil zwischen dem Umstand, daß da ein Schaf ist, und der Überzeugung, daß da ein Schaf ist, entgegen (4G) kein Kausalzusammenhang besteht. Nicht anders sieht es mit den ursprünglichen Gettier-Beispielen aus. In anderen Fällen sind die Intuitionen jedoch nicht so klar. Goldman diskutiert den Fall, daß ich eine Vase wahrnehme, die in der Tat vor mir auf dem Tisch steht. Man hat jedoch zwischen die Vase und meine Augen ein perfektes Hologramm dieser Vase projiziert – es ist erstaunlich, wie die moderne Technik die Phantasie der Philosophen beflügelt hat –, so daß mein Wahrnehmungseindruck von dem Hologramm und nicht von der Vase herrührt. Weiß ich dann, daß eine Vase vor mir steht? Es scheint, daß Bedingung (4G) nicht erfüllt ist; Goldman muß wohl sagen, daß ich es nicht weiß.

In Goldmans Rahmen muß ferner auch Wissen auf der Basis von Schlüssen erklärbar sein. Er präsentiert das folgende, etwas konstruierte Beispiel: Da ist ein Lavafeld auf Hawaii. Daß ein Vulkan ausgebrochen ist, ist eine bekannte Ursache dafür, daß auf dieser Insel sich nun ein Lavafeld befindet. Eine Person glaubt nun, daß da Lava herumliegt, hat gute Gründe für ihre Überzeugung etc. Also weiß sie, daß da Lava herumliegt. Sie schließt auf die Ursache zurück, daß auf der Insel früher ein Vulkan ausgebrochen ist. Goldman behauptet nun, daß diese Person seiner Theorie zufolge dann auch weiß, daß der Vulkan ausgebrochen ist. In der Tat ist (4G) erfüllt, weil der Ausbruch des Vulkans für ihre Überzeugung, daß der Vulkan ausgebrochen ist, indirekt ursächlich ist, nämlich via des Zwischeneffekts, daß da Lava herumliegt. Hätte jedoch einer zwischenzeitlich die Lava weggeräumt (zwecks Landschaftsverschönerung) und dann ein anderer wieder Lava ausgestreut (etwa als Filmkulisse), so wäre die Kausalkette in einer Weise unterbrochen, daß (4G) nicht mehr erfüllt ist. Ob diese beiden Geschichten von Goldman intuitiv plausibel sind oder nicht, muß jeder für sich selbst prüfen.

Es gibt noch weitere Beispiele, in denen Goldman darauf besteht, daß Zwischenglieder in Kausalketten unter bestimmten Bedingungen Wissen bewahren können, in anderen Fällen aber nicht. Betrachten wir die folgende Kausalkette, in der „→“ für ein Ursache-Wirkungs-Verhältnis und „--->„ für eine Relation des rationalen, plausiblen Erschließens stehen möge, welche Goldman ebenfalls als eine Art Kausalbeziehung deutet:

> $p \rightarrow y$ glaubt, daß $p \rightarrow y$ behauptet, daß $p \rightarrow x$ glaubt, daß y behauptet, daß p ---> x glaubt, daß y glaubt, daß p ---> x glaubt, daß p.

Bei einer solchen Schlußkette macht x einige Annahmen: etwa, daß, wenn jemand etwas behauptet, er es auch glaubt, und daß, wenn jemand etwas glaubt, es dann in der Regel auch so ist – daß also andere Personen im allgemeinen aufrichtig und verläßlich sind. Wenn diese Zusatzannahmen hinter den Pfeilen „--->“ zutreffen, dann weiß x im Falle der eben angegebenen Kausalkette, daß p. Aber wenn bei der nämlichen Kausalkette bestimmte Pannen passieren – etwa daß y sich beim Ausdruck seiner Überzeugung verspricht und versehentlich „nicht p“ sagt, x aber das „nicht“ in diesem Versprecher überhört und darum glaubt, daß p –, dann soll nach Goldman nicht der Fall sein, daß x weiß, daß p.

In den bisherigen Beispielen war die fragliche Überzeugung von dem Sachverhalt, den sie zum Inhalt hat, kausal abhängig. Goldman formuliert (4G) jedoch nicht mit Hilfe des Begriffs der kausalen Abhängigkeit, sondern mit Hilfe des schwächeren Begriffs des Kausalzusammenhangs; p und q hängen kausal zusammen gdw. es p_1,..., p_n gibt derart, daß $p = p_1$, $q = p_n$ und für alle $k = 1,..., n\text{-}1$ gilt, daß $p_{k\text{-}1}$ Ursache *oder* Wirkung von p_k ist. Goldman tut dies, weil er Wissen über zukünftige Ereignisse zulassen möchte. Sei q etwa der Sachverhalt, daß y morgen den Zustand p herbeiführen will. Dann könnte die folgende Kette auftreten:

$p \leftarrow q \rightarrow$ y sagt, daß q $\rightarrow$ x glaubt, daß y sagt, daß q --->
---> x glaubt, daß q ---> x glaubt, daß p

Hierin ist also q ursächlich für das morgige p; und insgesamt besteht zwischen p und „x glaubt, daß p“ ein Kausalzusammenhang im erläuterten Sinne.Gemäß (4G) weiß also x schon heute, daß p; so will es Goldman haben.

Man kann seine Intuitionen noch an anderen Beispielen testen. Ich habe gestern meinen Ofen angeheizt. Das führt normalerweise dazu, daß Rauch aus dem Kamin austritt. Es ist auch Rauch ausgetreten, aber ich habe das Haus nicht verlassen und es darum nicht gesehen. Weiß ich dann, daß Rauch aus dem Kamin ausgetreten ist? Ich weiß um die kausalen Zusammenhänge zwischen dem Anheizen und der Rauchentwicklung. Ich weiß aber auch, daß es immer wieder vorkommt, daß diese Kausalketten unterbrochen werden. Und ich habe keine Prüfung der Umstände vorgenommen, ob es nun tatsächlich geraucht hat oder nicht.

Wenn man, wie ich, in all diesen zweifelhaften Fällen der Meinung zuneigt, daß kein Wissen vorliegt, daß ich also z.B. nicht weiß, daß Rauch aus meinem Kamin kam, so wird man vielleicht (4G) dadurch verstärken wollen, daß man die Rede vom

Kausalzusammenhang durch die Rede von kausaler Abhängigkeit ersetzt. Dann resultiert allerdings ein sehr enger Wissensbegriff, wonach man allenfalls über die Vergangenheit wissen kann und rein gar nichts über die Zukunft. Denkt man die schon diskutierten Wahrnehmungen und Erinnerungen, so ist das korrekt; doch in voller Allgemeinheit mögen das andere wiederum intuitiv nicht plausibel finden. Ich will diese Frage jetzt nicht entscheiden.

So weit also zur externalistischen Behandlung des Wissensbegriffs. Schauen wir uns nun an, wie ein Internalist bei der Suche nach einer passenden Klausel (4) in der Analyse des Wissensbegriffs verfährt. Als Protagonist dient mir Keith Lehrer. Er möchte nicht auf äußere Faktoren zurückgreifen, sondern interne Kriterien dafür angeben, was gute Gründe ausmacht.[37] In den Gettier-Beispielen läuft seiner Ansicht nach folgendes schief: die Begründungen für die Überzeugungen sind entkräftbar („defeasible“) in dem Sinn, daß es wahre Aussagen gibt, die ihnen den Begründungsstatus entziehen würden, würde das Subjekt sie korrekt zur Kenntnis nehmen. Meine Begründung im Schaf-Beispiel war, daß dieser Hund wie ein Schaf aussieht. Würde man mich darüber aufklären, daß es sich um einen Bedlington-Terrier handelt, so würde die Begründung für meine Überzeugung und damit auch sie selbst entfallen. Würde man Schmid über die Ergebnisse seines Vorstellungsgesprächs aufklären, so fiele mit seiner Begründung auch die Überzeugung weg, daß derjenige, der die Stelle kriegt, 10 Mark in der Hosentasche hat. Nicht anders in den anderen Beispielen. Diese Überlegung führt Lehrer zu der folgenden Ergänzung der Analyse:

(4L) *x*' Begründung für *p* ist nicht entkräftbar.

Dabei gilt:

wenn *q* ein Grund von *x* für *p* ist, so entkräftet *r* den Grund *q* gdw. *r* wahr ist und (*q*&*r*) kein hinlänglicher Grund mehr für *p* ist.

Der hier benutzte Begründungsbegriff ist ersichtlich schwächer als der logische. Dort würde – technisch gesprochen, infolge der Monotonie logischer Folgerung – gelten, daß wenn *q p* begründet, dann auch jede Verstärkung von *q* weiterhin eine akzeptable Begründung liefert. Deduktive Gründe können also nie in Lehrers Sinn entkräftbar sein.

---

[37] In Lehrers und Paxsons Aufsatz in Bieri (1987) setzen sich die Autoren am Ende ihres Abschnitts IV kritisch mit Goldmans Theorie auseinander.

Doch ist, wie Lehrer selbst erklärt, (4L) immer noch zu stark. Die Bedingung ist hinreichend, aber nicht notwendig. Angenommen, ich weiß p. Nun ist es oft genug der Fall, daß ich irreführende wahre Informationen habe – irreführend in dem Sinn, daß sie nahelegen, daß p nicht der Fall ist. Unter solchen Umständen sind meine Gründe entkräftbar, und ich wüßte gemäß (4L) doch nichts. Das erscheint zu restriktiv. Lehrer erläutert das an seinem Grabski-Beispiel:

Mir ist Hans Grabski wohlbekannt. Eines Tages beobachte ich ihn dabei, wie er in der Unibibliothek Bielefeld ein Buch mitgehen läßt. Also sollte man doch meinen, ich wüßte, daß Hans Grabski ein Buch geklaut hat. Doch gibt es da noch Vater Grabski, der allen, die er erreichen kann, erzählt, daß sein Sohn Hans einen Zwillingsbruder Fritz hat, der für seine kleptomanischen Neigungen bekannt ist. Ich weiß von diesen Erzählungen von Vater Grabski nichts. Wüßte ich davon, würde ich doch sehr unsicher werden, ob es Hans Grabski war, der sich als Bücherdieb betätigt hat. Die Wahrheit freilich ist, daß Hans Grabski kleptoman ist, dies erheblich zur geistigen Zerrüttung seines Vaters beigetragen hat, der nun einen zweiten Sohn erfunden hat und alle Untaten seines einzigen Sprößlings auf diesen schiebt. Die Erzählungen von Vater Grabski entkräften in Lehrers Sinn meine Gründe für die Überzeugung, Hans Grabski stehle Bücher. Doch die bloße Existenz solcher potentiellen Entkräftungen, meint Lehrer, untergräbt noch nicht meinen Wissensanspruch; in der Grabski-Geschichte weiß ich, daß Hans Grabski ein Buch geklaut hat – egal, was der Vater erzählt. Darum schlägt Lehrer vor, die Definition des Begriffs der Entkräftbarkeit in der folgenden Weise abzuändern:

> Wenn q ein Grund von x für p ist, so entkräftet r den Grund q gdw. r wahr ist, x hinlängliche Gründe dafür hat zu glauben, daß r falsch ist, und $(q\&r)$ kein hinlänglicher Grund für p ist.

Auf das Grabski-Beispiel angewandt: Wenn ich glaube, daß Hans Grabski Bücher klaut, so habe ich dafür meine guten Gründe, ohne daß dabei eine Überzeugung über Vater Grabski eine Rolle spielte. Also habe ich auch keine hinlänglichen Gründe für eine Überzeugung des Inhalts, daß Vater Grabski solche Geschichten erzählt. Und damit entkräftet die Tatsache, daß Vater Grabski solche Geschichten erzählt, gemäß Lehrers verbesserter Definition nicht meine Begründung für meine Einschätzung von Hans Grabski. Um dies mit dem Ferrari-Beispiel (beschränkt auf die Herren Protz und Krämer) zu vergleichen: da ist r „Protz hat keinen Ferrari“. Diese Aussage ist wichtig, denn meine guten Gründe für die Überzeugung, einer im Hörsaal besitze einen Ferrari, stützen sich darauf, daß ich „Protz hat keinen Ferrari" für falsch halte; ich habe ja gerade gute Gründe für das Gegenteil. Mithin ist Lehrers verbesserte Definition von „entkräftbar" in diesem Falle anwendbar.

Die internalistische Analyse erweist sich bis hierhin schon als reichlich kompliziert. Allerdings ist Lehrer mit der von ihm erreichten Analyse noch immer nicht zufrieden; weitere Komplikationen sind nötig, welche die geneigte Leserin aber bei Lehrer selbst studieren möge. Damit will ich es auch mit meinen inhaltlichen Ausführungen zur Wissensanalyse bewenden lassen. Die zwei Hauptvertreter der zwei hauptsächlichen Positionen dazu, worin die gesuchte vierte Wissensbedingung bestehen könnte, habe ich vorgestellt. Die Kluft zwischen ihnen scheint groß; doch manche machen sich anheischig, unter gewissen Voraussetzungen ihre Äquivalenz nachzuweisen.[38] Die Diskussion könnte auch auf andere Weise überflüssig sein; manche meinen jedenfalls, daß schon die dritte und damit erst recht jene ominöse vierte Wissensbedingung nicht benötigt würde, daß also Wissen einfach wahrer Glauben sei. Das Argument dafür ist, grob gesagt, daß Überzeugungen, rationaliter jedenfalls, ipso facto begründete seien.[39] Altmeister David Lewis macht sich für den sogenannten „relevant alternatives"-Ansatz stark.[40] Und so weiter.

Das ist der ganz normale Effekt philosophischer Analyse. Einerseits bedeuten all die Differenzierungen und Detaildiskussionen einen großen Zugewinn an Klarheit und Einsicht. Andererseits wird die Meinungsbildung dadurch enorm erschwert; man wird sich eher immer unsicherer, was man schließlich für wahr und richtig halten soll. Jedenfalls gibt es keine allgemein akzeptierte Lehre, jeder muß sich selbst seine Meinung bilden. Klar geworden ist bei alledem auch, daß der Glaubensbegriff erkenntnistheoretisch vorgeordnet ist. Diesem müssen wir uns nun endlich zuwenden.

---

[38] Z.B. Hunter (im Erscheinen).

[39] So ganz dezidiert von Kutschera (1982), Kapitel 1. Auf Englisch vertritt das Sartwell 1991) und (1992)

[40] Lewis (im Erscheinen).

4. Statische Glaubenstheorie

Wie die meisten wissenschaftlichen Theorien setzt sich auch die Glaubenstheorie aus einem statischen und einem dynamischen Teil zusammen. Im statischen Teil geht es immer darum, den Zustand des jeweiligen wissenschaftlichen Gegenstandes für einen gegebenen Zeitpunkt sowie die Gesetze, denen er zu diesem Zeitpunkt unterliegt, zu beschreiben. Darauf aufbauend geht es dann im dynamischen Teil darum, die zeitliche Entwicklung und Veränderung des Zustands dieses Gegenstands und die Gesetze dieser Veränderung zu erfassen. Daher will ich hier zunächst nur um die statische Glaubenstheorie behandeln.

Die Frage ist nun also, wie sich der doxastische Zustand eines Erkenntnissubjekts zu einem gegebenen Zeitpunkt geeignet umfassend beschreiben läßt.[41] Dazu gibt es zwei hauptsächliche Möglichkeiten: die eine verfährt qualitativ, die andere probabilistisch. Im Rahmen einer probabilistischen Beschreibung werden, in der Gestalt von Wahrscheinlichkeiten, Glaubensgrade spezifiziert. Das ermöglicht eine entsprechend feine Beschreibung. Bei einer qualitativen Herangehensweise begnügt man sich mit einer gröberen Einteilung; es gibt da nur drei Fälle: entweder glaubt x, daß p; oder er glaubt, daß non-p; oder er ist neutral eingestellt, d.h., er glaubt nicht, daß p, und er glaubt auch nicht, daß non-p. Die beiden ersten Fälle werden wir im folgenden, unter Weglassung des Hinweises auf das Erkenntnissubjekt, mit $G(p)$ bzw. $G(\text{non-}p)$ mitteilen.[42]

Angesichts der Tatsache, daß die qualitative Beschreibung doxastischer Zustände ersichtlich so viel gröber ausfällt als eine probabilistische, fragt sich, welches Recht diese Betrachtungsweise eigentlich besitzt. Drei Gründe lassen sich zur Beantwortung dieser Frage anführen. Erstens benutzen wir im Alltag ständig die qualitative Beschreibungsweise doxastischer Zustände; auch in gewissen theoretischen Zusammen-

---

[41] Terminologische Anmerkung: In der Literatur wird oft von *epistemischen* Zuständen gesprochen. Von der Ableitung aus dem Griechischen her ist das nicht korrekt, denn danach sind epistemische Zustände solche, die *Wissen* betreffen. Um über *Überzeugungen* oder *Glauben* zu reden, ist das Adjektiv „*doxastisch*“ korrekt.

[42] An dieser Stelle wieder eine Mahnung zur Vorsicht: bei Spezifizierungen von doxastischen Zuständen muß man auf verschiedene Arten der Negation achten, die in der Umgangssprache nicht immer zureichend unterschieden werden. Es ist ein Unterschied, ob man sagt: „x glaubt, daß nicht p“ (innere Negation) oder „x glaubt nicht, daß p“ (äußere Negation). Nur der zweite Fall ist damit verträglich, daß x sich in Bezug auf p einer Meinung enthält.

hängen scheint sie unentbehrlich. Zum zweiten ist sie zwar grob, aber doch nicht unzulässig grob. Der wichtigste Grund dafür, sie zu benutzen, liegt freilich darin, daß entgegen dem ersten Anschein gar nicht klar ist, wie sich die qualitative Beschreibung in feinere Beschreibungen einpassen oder gar auf solche reduzieren läßt. Das werde ich weiter unten noch genauer ausführen. Und selbst wenn eine derartige Reduktion doch möglich sein sollte, wäre es gut zu verstehen, was da eigentlich reduziert werden soll. Jedenfalls will ich zunächst die qualitative Beschreibungsweise diskutieren.

Worin könnten die statischen Gesetzen der Glaubenstheorie bestehen? Zu einem gegebenen Zeitpunkt hat ein Erkenntnissubjekt sehr, sehr viele Überzeugungen. Es ist ja etwas nicht nur dann meine Überzeugung, wenn es mir gerade durch den Kopf geht; die meisten meiner Überzeugungen habe ich lediglich in dispositionaler Form; sie schlummern gewissermaßen in mir und ließen sich, wenn nötig, aktualisieren. Die Gesetze der statischen Glaubenstheorie schreiben nun gewisse Bedingungen fest, denen die Menge unserer Überzeugungen zu einem gegebenen Zeitpunkt unterliegen muß. Es sind sogenannte Koexistenzgesetze, nicht anders als z.B. das Boyle-Mariotte'sche Gesetz in der Physik, das Druck, Volumen und Temperatur eines idealen Gases zu einem gegebenen Zeitpunkt miteinander in Beziehung setzt, indem es aussagt, daß $p \cdot V/T$ = konstant, wobei p für Druck, V für Volumen und T für die Temperatur des Gases steht. In der statischen Glaubenstheorie suchen wir vergleichbare Gesetze für das Glaubensprädikat G.

Gleich als erstes erhebt sich da die schwierige Frage, wofür denn das „p“ in $G(p)$ steht. Im Kapitel 7 wird das Gegenstand ausführlicher Diskussionen sein, aber es ist unvermeidlich, schon an dieser Stelle vorläufig darauf einzugehen. Unter Philosophen hat sich in den letzten Jahrzehnten die Rede von *propositionalen Einstellungen* eingebürgert. Zu glauben, daß p, aber auch zu wünschen, zu hoffen, zu fürchten, zu beabsichtigen, daß p, etc., all dies fällt unter diesen Begriff. In der Tat verwenden viele diesen Begriff so weit, daß alle psychologische Verben, die daß-Sätze als Komplement nehmen, Ausdrücke für propositionalen Einstellungen sind. Andere verstehen ihn strikter dahingehend, daß ein Verb nur dann für eine propositionale Einstellung steht, wenn es als Relation zu Propositionen zu interpretieren ist.[43] Daran schließt

[43] Eine andere Ambiguität: Manche Philosophen bezeichnen als propositionale Einstellungen die konkreten geistigen Zustände eines Individuums, über die mit „glauben“, „wünschen“, etc. gesprochen wird – unabhängig davon, wie diese Zustände genau zu verstehen sind; andere bezeichnen Einstellungen, also das Glauben, Wünschen etc. selbst als propositionale Einstellungen – unabhängig davon, wie sie sich im Subjekt genau realisieren.

sich gleich die Frage an, was *Propositionen* sein sollen. Offenbar sind sie etwas, was von Sätzen ausgedrückt wird; aber die Frage ist eben, wie das Ausgedrückte zu verstehen ist. In der Tat ist der Begriff der Proposition eine der schwierigsten in der Philosophie; ich werde im folgenden nur eine einfache Charakterisierung geben, die viele (auch ich) im Prinzip für angemessen halten und die wir jedenfalls diesem und den nächsten Kapiteln zugrunde legen können:

Sätze, und daher damit auch die Komplement- oder Inhaltssätze von Einstellungsverben, weisen einen bestimmten Informationsgehalt auf. Dieser läßt sich über den Ausschluß von Möglichkeiten charakterisieren: Mit Ω bezeichne ich den sogenannten Möglichkeitsraum, die Menge all der für relevant erachteten Möglichkeiten. Der kann klein oder groß sein, je nachdem, welches Problem behandelt werden soll. Wenn wir mögliche Ziehungsergebnisse beim Lotto betrachten wollen, dann ist der Möglichkeitsraum die Kombination aller möglichen Ziehungen von 6 Zahlen plus einer Zusatzzahl. Offenbar ist das ein vergleichsweise kleiner Möglichkeitsraum. Ein anderes Beispiel: alle möglichen Verteilungen von Prozentzahlen auf die 13 Listen, die zur NRW-Landtagswahl 1995 angetreten sind. Wenn da einer vor der Wahl gesagt hat: die SPD erhält 50% oder mehr, so schließt diese Äußerung für ihre Wahrheit alle Verteilungen der Prozente aus, in denen die SPD weniger als 50% erhält. Wenn einer prognostiziert hätte, daß die SPD exakt 50% erhält, so hätte er damit noch mehr Möglichkeiten ausgeschlossen.

Einer Zunahme des Informationsgehaltes einer Aussage entspricht also eine Zunahme der ausgeschlossenen Möglichkeiten. Das erlaubt, die Information über Teilmengen des Möglichkeitsraums zu charakterisieren. Ich werde nun fürs weitere eine *Proposition* einfach mit einer Teilmenge von Ω identifizieren, eben mit der Menge von Möglichkeiten, in denen sie wahr ist; wenn es nötig ist, die Relativierung auf Ω explizit zu machen, rede ich auch von Propositionen *über* Ω. Ω selbst ist eine solche Teilmenge; es ist die notwendige oder tautologische Proposition – diejenige, die immer wahr ist und keinen Informationsgehalt hat, weil sie keine Möglichkeiten ausschließt. Die leere Menge Ø ist ebenfalls eine Teilmenge von Ω; sie ist die unmögliche oder kontradiktorische Proposition.[44] Alle informationshaltigen Propositionen liegen im Möglichkeitsraum zwischen diesen beiden Extremen. Die Inhalte von Überzeugungen sind also nach dem im weiteren benutzten Bild Propositionen, charakterisiert als Teilmengen des Möglichkeitsraumes.

---

[44] Es ist korrekt, bei der Charakterisierung von Ω und Ø den bestimmten Artikel zu benutzen; beide Objekte sind einzig.

Kehren wir nach dieser Klärung zu der Frage nach den Gesetzen der statischen Glaubenstheorie (in der qualitativen Beschreibungsweise) zurück. Diese werden in der sogenannten doxastischen oder epistemischen Logik diskutiert; es werden dort vier Standardgesetze formuliert:[45]

(G1) Wenn $G(p)$, dann nicht $G(\neg p)$.

(G2) Wenn $G(p)$ und $G(q)$, dann $G(p \wedge q)$.

(G3) Wenn $G(p \wedge q)$, dann $G(p)$ (und natürlich dann auch $G(q)$).

(G4) Wenn p und q logisch äquivalent sind, dann $G(p) \leftrightarrow G(q)$.

Das erste Gesetz besagt, daß rationale Überzeugungen konsistent sind, das zweite und das dritte behaupten die Abgeschlossenheit rationalen Glaubens unter Konjunktionsbildung und unter Konjunktionsauflösung. Das vierte Gesetz schließlich ermöglicht die Substitution logisch äquivalenter Propositionen füreinander auch in Überzeugungskontexten. Wie wir sehen werden, ist dieses vierte Gesetz im Vergleich zu den anderen äußerst stark; wenn sich Kritik gegen diese Theorie richtet, wird sie fast immer an diesem Gesetz einhaken.

Aus diesen Gesetzen folgt, daß die Menge der Überzeugungen in einem doxastischen Zustand logisch konsistent und deduktiv abgeschlossen ist. Das sieht man folgendermaßen: Das erste Gesetz schließt zunächst nur unmittelbare logische Widersprüche in den Überzeugungen aus. Aber mit den anderen Gesetzen sind auch „verstecktere" Widersprüche ausgeschaltet: Gelte $G(p_i)$ für alle $i = 1,...,n$, und seien $p_1,...,$ p_n widersprüchlich. Es gilt dann, infolge der Abgeschlossenheit unter Konjunktionsbildung (Gesetz G2), $G(p_1 \wedge ... \wedge p_{n-1})$ und zugleich $G(p_n)$. Ferner folgt aus $p_1 \wedge ... \wedge p_{n-1}$ logisch, daß $\neg p_n$ (das ist eine Möglichkeit, den Widerspruch zum Ausdruck zu bringen). Das bedeutet, daß es einen zu $p_1 \wedge ... \wedge p_{n-1}$ logisch äquivalenten Satz der Form $q \wedge \neg p_n$ gibt. Mit dem Gesetz (G4) folgt also weiter, daß $G(q \wedge \neg p_n)$. Aus letzterem ergibt sich schließlich mit dem Gesetz (G3) $G(\neg p_n)$, was in direktem Widerspruch zu $G(p_n)$ steht, d.h. dem Gesetz (1) zuwiderläuft. Mithin schließen die Gesetze die angenommene Konstellation aus.

---

[45] Literatur zur epistemischen Logik: Hintikka (1962); von Kutschera (1976); Lenzen (1980). – Anmerkung zum Zeichengebrauch: ich verwende die üblichen Symbole für die logischen Verknüpfungen: „$\neg$" für „nicht", „$\wedge$" für „und", „$\vee$" für das einschließende „oder", „$\rightarrow$" für die materiale Implikation und „$\leftrightarrow$" für die materiale Äquivalenz.

Die deduktive Abgeschlossenheit eines Glaubenssystems ergibt sich so: Folge aus $p_1,...,p_n$ logisch, daß q. Dann gibt es wiederum einen zu $p_1 \wedge ... \wedge p_n$ logisch äquivalenten Satz der Form p∧q. Aus $G(p_1)$,..., $G(p_n)$ folgt nun mit Gesetz (G2) zunächst $G(p_1 \wedge ... \wedge p_n)$. Daraus folgt weiter mit der Substituierbarkeit logischer Äquivalenzen (Gesetz G4), daß $G(p \wedge q)$. Und daraus ergibt sich mit Gesetz (G3) schließlich $G(q)$. Insgesamt bedeutetdas, daß ein Subjekt die logischen Folgerungen aus seinen Überzeugungen ebenfalls glaubt.

Wo kommen diese Gesetze der qualitativen statischen Glaubenstheorie her, was ist ihr Status? Es sieht nicht so aus, als hätten sie unmittelbar empirischen Gehalt; denn es passiert ja zuweilen, daß wir entdecken, daß unser Überzeugungssystem Widersprüche enthält. Also scheint es, daß diese Gesetze normativen Status aufweisen; sie charakterisieren, was ein rationales Überzeugungssystem ausmacht. So jedenfalls lautet die Standardantwort auf die eben gestellte, drängende Frage.

Der Standardantwort wird ein Standardeinwand entgegengehalten: Rationalitätsgesetze müssen menschenmöglich sein in dem Sinne, daß normale Individuen sie auch erfüllen können. Aber die angegebenen Gesetze stellen übermenschliche Forderungen. Nicht nur, daß wir de facto immer wieder Widersprüche in unseren Überzeugungen antreffen und auf keinen Fall Überzeugungen über alle deduktiven Konsequenzen auch nur derjenigen Teilmenge unserer Überzeugungen haben, die uns unmittelbar vor Augen stehen – schlimmer noch, der Begriff des logischen Widerspruchs ist gar nicht entscheidbar. (Das bedeutet, daß es kein mechanisches Verfahren gibt, dessen Anwendung eine Überprüfung aller Mengen von Sätzen auf ihre Konsistenz hin erlaubt, in dem Sinne, daß dieses Verfahren uns in endlich vielen Schritten sagt, ob eine Menge konsistent ist oder nicht.) Selbst den klügsten Leuten passiert es, daß sie widersprüchliche Überzeugungen haben. Ein Beispiel ist Gottlob Frege, der Urvater der modernen Logik, der es als erster den Versuch unternahm, die Mathematik mit einem System der axiomatischen Mengentheorie zu unterfüttern. Doch, ach, das System war inkonsistent.[46]

Wer die qualitativen statischen Glaubensgesetze so kritisiert, wird zugeben, daß es irrational ist, wenn jemand ihm bekannte logische Folgerungen nicht doxastisch umsetzt. Aber es ist nicht irrational, nicht alle logischen Konsequenzen zu überblicken.

[46] Cf. Frege (1893). Es war Bertrand Russell, der dieses System durch den Hinweis auf etwas abschoß, das heute als Russellsche Antinomie bekannt ist: die Menge aller Mengen, die sich selbst nicht als Element enthalten. Diese Menge enthält sich selbst genau dann als Element, wenn sie sich selbst nicht als Element enthält – was ein Widerspruch ist. In Freges System gibt es diese Menge.

Dieser Kritik läßt sich jedoch entgegnen, daß sie genau genommen viel zu weit geht. Wenn man sich ihre Position zu eigen macht, ist nämlich nicht mehr zu sehen, welche Gesetze für Überzeugungssysteme überhaupt noch übrig bleiben; es gibt eigentlich keine vernünftige Abschwächung. Außerdem bleibt immer noch die Möglichkeit, die Gesetze wenigstens als Idealisierungen aufzufassen, denen normal rationale Individuen immerhin annähernd genügen. Das reicht vielleicht, um sie als Gesetze für den statischen Glaubensbegriff gebrauchen zu können.

In der Tat ist es so, daß sich die Auseinandersetzung eigentlich nur um das Gesetz (G4), die Substituierbarkeit logischer Äquivalenzen, dreht. Alle für verdächtig angesehenen Folgerungen ergaben sich nur unter Einsatz dieses Gesetzes. Doch – und das ist nun der entscheidende Punkt – wenn es wirklich Propositionen, also Teilmengen des Möglichkeitsraumes, sind, die die Gegenstände des Glaubens bilden, so ist dieses Gesetz unangreifbar. Denn wenn p die Proposition A, also die Menge A von Möglichkeiten, ausdrückt und q mit p logisch äquivalent ist, so drückt q genau dieselbe Proposition A aus, hat also denselben Informationsgehalt wie p und schließt dieselben Möglichkeiten aus. Das Gesetz (G4) ist also nichts weiter als eine Folge unserer Charakterisierung von Propositionen und der Tatsache, daß wir Überzeugungen als propositionale Einstellungen im engeren Sinne behandeln. Diese Grundannahme fiele mit dem Gesetz (G4). Aber dann müßte man sich nach einem ganz anderen Theorieformat umsehen. Wie dem auch sei, offensichtlich müßten wir nun genauer unter die Lupe nehmen, wie Überzeugungsinhalte genau aufzufassen sind, was wir aber erst im Kapitel 7 tun werden.

Sind die erwähnten vier Gesetze der qualitativen statischen Glaubenstheorie alles, was die Theorie anzubieten hat? Es gibt noch mehr, nämlich sogenannte Reflexionsprinzipien:

(G5) Wenn $G(p)$, dann auch $G(G(p))$.

(G6) Wenn $G(G(p))$, dann auch $G(p)$.

Wenn ich glaube, daß p, dann glaube ich gemäß Gesetz (G5) auch, daß ich glaube, daß p – und so weiter zu immer tieferen Verschachtelungen von Glaubenseinstellungen. Das sieht auch wieder wie eine ziemliche Idealisierung aus, angesichts derer die eben geführte Diskussion einschlägig ist. Das Gesetz (G6) hingegen scheint unproblematisch.

Diese Reflexionsprinzipien ermöglichen – um ein Beispiel für ihre Tragweite zu geben – eine schöne Behandlung von Moores Paradox. Dahinter verbirgt sich folgendes: Es ist klarerweise möglich, daß ich glaube, daß non-p, aber p der Fall ist. Ich glaube, daß es nicht regnet, aber nichts spricht gegen die Möglichkeit, daß es trotzdem regnet. Doch ist es Unsinn, wenn ich sage – und damit meiner Überzeugung Ausdruck verleihe –, daß es regne, ich aber nicht glaube, daß es regne. Mit Hilfe der Reflexionsprinzipien läßt sich erläutern, inwiefern das Unsinn ist. Formal läßt sich Moores paradoxaler Satz so darstellen: $G(p \wedge G(\neg p))$. Daraus folgt mit dem Gesetz (G3), daß $G(G(\neg p))$, und daraus mit dem Reflexionsprinzip (G6), daß $G(\neg p)$. Doch folgt mit dem Gesetz (G3) aus Moores Satz auch $G(p)$ – womit der Widerspruch zutage liegt. Meine Äußerung ist also in dem Sinne unsinnig, daß sie eine gemäß den doxastischen Gesetzen inkonsistente Überzeugung ausdrückt.

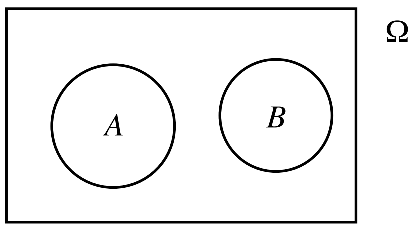
So weit meine sehr kurze Darstellung der Herangehensweise der epistemischen Logik an die statische Glaubenstheorie. Wenden wir uns nun der probabilistischen Version der statischen Glaubenstheorie zu. Auch diese Art der Beschreibung ist uns aus alltäglichen Wendungen wohlvertraut: wir sagen, daß es *wahrscheinlich* auch morgen nicht regnen wird, daß es *eher der Fall* ist, daß Werder Deutscher Meister wird als Dortmund[47], daß *mit an Sicherheit grenzender Wahrscheinlichkeit* es kein Erdbeben in diesem Jahr die Uni Bielefeld in Trümmer legt, usw. Diese Redeweisen gilt es nun innerhalb eines präzisen theoretischen Rahmens zu explizieren. Dafür gibt es mittlerweile verschiedenartige Möglichkeiten, auf die ich später noch ein paar Hinweise geben werde; der bei weitem gebräuchlichste und best ausgebaute Rahmen ist aber der der Wahrscheinlichkeitstheorie, und nur ihn will ich im folgenden etwas ausführen. Es ist dafür unerläßlich, ein wenig ins technische Detail zu gehen.

Der zentrale Begriff ist der des *Wahrscheinlichkeitsmaßes*:

> P ist ein Wahrscheinlichkeitsmaß für Ω gdw. P eine Funktion ist, die jeder Proposition über Ω eine reelle Zahl zuordnet, so daß gilt:
>
> (P1) $P(A) \geq 0$ (Nicht-Negativität),
>
> (P2) $P(\Omega) = 1$ (Normiertheit),
>
> (P3) Wenn A und B einander ausschließen, dann ist $P(A \text{ oder } B) = P(A) + P(B)$ (Additivität).

[47] Wie man sieht, kann man sich bei solchen Beurteilungen auch irren. Als der Text im Mai 95 vorgetragen und geschrieben wurde, war Werder noch vorn; drei Wochen später feierte Dortmund die Meisterschaft. Ich hoffe, daß ich mit dem folgenden Beispiel nicht genauso daneben liege.

Man kann sich das anhand von Venn-Diagrammen gut veranschaulichen:



Ω ist das Rechteck, welches für den ganzen Möglichkeitsraum stehe. *A* und *B* sind zwei Propositionen; daß sie sich nicht überlappen, bedeutet, daß sie einander logisch ausschließen. In diesem Fall, so sagt das Additivitätsaxiom (P3), kann man die Wahrscheinlichkeiten von *A* und *B* einfach addieren, um die Wahrscheinlichkeit dafür zu erhalten, daß eine der beiden Propositionen wahr ist. Mit Flächeninhalten kann man offenbar ebenso operieren. In der Tat ist das Flächeninhaltsmaß für Teilflächen einer auf 1 normierten Gesamtfläche ein Wahrscheinlichkeitsmaß. Man spricht hier von der geometrischen Interpretation des Wahrscheinlichkeitsbegriffs, die für die Visualisierung wahrscheinlichkeitstheoretischer Sachverhalte ungemein nützlich ist.

Aus den drei Axiomen ergeben sich sofort einige elementare Folgerungen. Zunächst folgt:

(P4) $\quad P(\emptyset) = 0.$

Denn $P(\Omega$ oder $\emptyset)$ erfüllt die Annahme des Axioms (P3); damit ist $P(\Omega$ oder $\emptyset) = P(\Omega) + P(\emptyset)$. Da aber $P(\Omega$ oder $\emptyset)$ nichts anderes ist als $P(\Omega)$, was nach Axiom (P2) gleich 1 ist, erhält man $1 = 1 + P(\emptyset)$, d.h. (4).

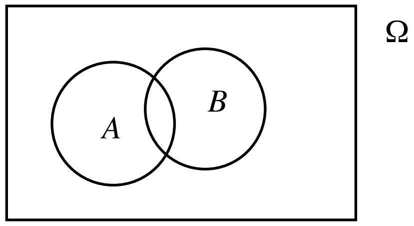
(P5) $\quad P(\text{non-}A) = 1 - P(A).$

Denn $P(A$ oder non-$A)$ erfüllt die Voraussetzung von (P3); daher gilt $P(A$ oder non-$A) = P(A) + P(\text{non-}A)$. Da aber ($A$ oder non-A) dasselbe ist wie Ω, gilt mit (P2) $P(A) + P(\text{non-}A) = 1$, d.h. (P5).

Man fragt sich vielleicht, ob es auch eine Version des Additivitätsprinzips gibt, die im Gegensatz zu (P3) ohne die Voraussetzung auskommt, daß A und B einander ausschließen. Ja; sie heißt das verallgemeinerte Additivitätsprinzip:

(P6) $P(A \text{ oder } B) = P(A) + P(B) - P(A \text{ und } B)$.

Statt eines Beweises sei zur Illustration wieder ein Venn-Diagramm dargeboten:



Die Kreise symbolisieren die Propositionen A und B. Wenn man aus ihren Wahrscheinlichkeiten (Flächen) schlicht die Summe bilden würde, würden aber alle die Elemente, die A und B gemeinsam haben, doppelt gezählt werden; darum muß von der Summe noch die Wahrscheinlichkeit der Linse, die A und B gemeinsam haben und die die Proposition (A und B) repräsentiert, abgezogen werden.

(P7) Wenn $\Omega = \{w_1,...,w_n\}$ und $P(\{w_i\}) = x_i$, so ist $x_i \geq 0$ für jedes $i = 1, ...,n$ und $x_1 + ... + x_n = 1$; und es gilt für jedes A:
$$P(A) = \sum_{w_i \in A} x_i \,.$$

Das ist ebenfalls eine einfache Folgerung; insbesondere ergibt sich die zuletzt genannte Eigenschaft einfach durch iterierte Anwendung der Additivität. Die Folgerung ist aber wichtig; sie bedeutet, daß man, zumindest wenn der Möglichkeitsraum Ω endlich ist, die Wahrscheinlichkeiten beliebiger Propositionen auf die Wahrscheinlichkeiten der einzelnen möglichen Fälle zurückspielen kann; zur Festlegung eines Wahrscheinlichkeitsmaßes reicht es also schon hin, die Zahlen x_1,..., x_n in der angegebenen Weise zu spezifizieren.

Diese letzte Folgerung läßt sich anhand einer anderen wichtigen, nämlich der sogenannten klassischen Interpretation des Wahrscheinlichkeitsbegriffs erläutern. Das ist jene Konzeption, die man anfänglich zur Erhellung von Problemen mit Glücksspielen benutzt hat. Z.B. hat man es mit zwei Würfeln zu tun, die zusammen geworfen werden. Der Möglichkeitsraum besteht also aus allen möglichen Paaren von Augenzahlen, d.h. aus insgesamt 36 Elementen. Diese werden unter der klassischen Wahrscheinlichkeitskonzeption alle als gleich wahrscheinlich betrachtet. Damit läßt sich die Antwort auf die Frage nach der Wahrscheinlichkeit einer beliebigen Proposition *A*über dem Möglichkeitsraum ganz einfach geben:

$$P(A) = \frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl der möglichen Fälle}} \,.$$

Fragt man sich also, was die Wahrscheinlichkeit dafür ist, mit zwei Würfeln 8 Augen zu erzielen, kommt man auf 5/36.

Ein weiterer wichtiger Begriff ist der der *bedingten Wahrscheinlichkeit*. Sie sitzen gerade vor dem Fernseher und verfolgen die Ziehung der Lottozahlen. Die Wahrscheinlichkeit, daß Ihre Tippreihe 6 Richtige enthält, ist minimal größer als 1:14 Millionen (genau 1:13.983.816). Die ersten drei Zahlen, die gezogen wurden, sind erstaunlicherweise schon richtig. Da, oh Schreck, verursacht ein Stromausfall das Ende der Übertragung. Nun sind Sie doch sicherlich neugierig zu wissen, wie groß die Wahrscheinlichkeit für den erwünschten Sechser jetzt ist – gegeben, Sie haben schon drei Richtige. Leider immer noch nicht sehr groß, nämlich 1:15.180. Wie rechnet man das aus? Dazu benötigt man bedingte Wahrscheinlichkeiten. Dazu wird eine zweistellige Funktion $P(-\mid-)$ wie folgt definiert:

Wenn $P(A) \neq 0$, so ist die *durch A bedingte Wahrscheinlichkeit von B* erklärt als

$$P(B \mid A) = \frac{P(A \text{ und } B)}{P(A)} \,.$$

Damit läßt sich Ihre Neugier befriedigen. Sie wollten wissen, was *P*(6 Richtige | die ersten drei sind richtig) ist. Definitionsgemäß ist das *P*(6 Richtige) / P(die ersten drei sind richtig), da ja *P*(6 Richtige) gleich *P*(6 Richtige und 3 Richtige) ist; und wenn man das berechnet, ergibt sich die obige Zahl. Im Fall der klassischen Wahrscheinlichkeit, die ja auch auf das Lotto-Beispiel paßt, vereinfacht sich die Definition offenbar noch zu der Formel

$$P(B \mid A) = \frac{\text{Anzahl der für } A \text{ und } B \text{ günstigen Fälle}}{\text{Anzahl der für } A \text{ günstigen Fälle}} \,.$$

Denn setzt man die Erklärung der klassischen Wahrscheinlichkeit in die Definition der bedingten ein, so kürzt sich im Nenner und im Zähler die Anzahl der insgesamt möglichen Fälle weg.

Aus der Definition der bedingten Wahrscheinlichkeit ergeben sich sofort zwei weitere, oft verwandte Theoreme:

(P8) $P(A \text{ und } B) = P(A) \cdot P(B \mid A)$ (Multiplikationsprinzip),

(P9) $P(A_1 \text{ und } A_2 \text{ und } \ldots A_n) =$
$P(A_1) \cdot P(A_2 \mid A_1) \cdot \ldots \cdot P(A_n \mid A_1 \text{ und } A_2 \text{ und } \ldots A_{n\text{-}1})$
(verallgemeinertes Multiplikationsprinzip).

Was hier nun steht, die Axiome (P1)-(P3) und ihre Folgerungen (P4)-(P9), sind alles statische Gesetze für Glaubensgrade. Doch wieso sollte man gerade diese Axiome für den Begriff des graduellen Überzeugtseins akzeptieren? Das ist nicht ohne weiteres klar. Ich erwähnte schon, daß diese Axiome für die geometrische Interpretation und ebenso für die klassische Auffassung des Wahrscheinlichkeitsbegriffs zutreffen. Es gibt noch eine weitere unstrittige Interpretation der Wahrscheinlichkeitsaxiome: nämlich als relative Häufigkeiten; auch diese verhalten sich gemäß (P1)-(P3). Die relative Häufigkeit eines Merkmals A in einer Grundgesamtheit Ω bestimmt sich genau nach der Formel für die klassische Wahrscheinlichkeit, da ja für die relative Häufigkeit jeder Fall eo ipso gleich viel zählt. Es gibt auch bedingte relative Häufigkeiten. Betrachten wir etwa die Landtagswahl 1995 in NRW. Was ist die relative Häufigkeit der CDU-Wähler in NRW? Die Grundgesamtheit ist die Menge aller Wahlberechtigten. Die Wahlbeteiligung, d.h. die relative Häufigkeit der Wählenden unter den Wahlberechtigten, lag bei 64%. Die CDU erzielte ein Wahlergebnis von knapp 38%; das ist die bedingte relative Häufigkeit der CDU-Wähler unter den Wählenden. Und damit ergibt sich gemäß (P8), daß die relative Häufigkeit der CDU-Wähler in NRW bei etwas über 24% liegt.

Doch auch wenn Flächen, Fallzahlen und relative Häufigkeiten die Wahrscheinlichkeitsaxiome erfüllen – was zweifelsohne unser Verständnis dieser Axiome fördert –, so sind sie nicht eigentlich Wahrscheinlichkeiten. Flächen dienen der geometrischen Veranschaulichung; Fallzahlen werden nur dadurch zu Wahrscheinlichkeiten, daß man alle Fälle als gleich wahrscheinlich betrachtet und so den Wahrscheinlichkeitsbegriff allererst ins Spiel bringt; und relative Häufigkeiten hängen über das Gesetz der großen Zahlen sehr eng mit Wahrscheinlichkeiten zusammen, sind aber selbst keine. Was sind Wahrscheinlichkeiten dann? Man kann Wahrscheinlichkeiten subjektiv interpretieren, als Glaubensgrade eines Subjekts zu einem bestimmten Zeitpunkt; man kann Wahrscheinlichkeiten objektiv interpretieren, vielleicht als Grenzwerte beobachteter relativer Häufigkeiten oder auch als numerisch-graduell spezifizierbare Geneigtheiten von Gegenständen, die und die Eigenschaft oder das und das Verhalten aufzuweisen; und so weiter. In der Tat liegt in der Interpretation des Wahrscheinlichkeitsbegriffs ein erhebliches philosophisches Problem. Die gerade gegebenen groben Charakterisierungen sind ja alles andere als klar; und ob sie dann auch angemessen sind, ist erst recht nicht klar. Und so existieren, seit der Begriff in der intellektuellen Welt eine Rolle spielt, nebeneinander mehr oder weniger objektive und subjektive Deutungen, deren schwieriges Verhältnis zueinander Gegenstand forwährender Diskussion ist.[48]

Für unsere erkenntnis- oder, genauer, glaubenstheoretischen Zwecke ist eindeutig die subjektive Interpretation von Wahrscheinlichkeiten als Glaubensgrade eines Subjekts die einschlägige. Doch stellt sich dann die umgekehrte Frage: Warum sollen sich die subjektiven Glaubensgrade eines Subjekts im Sinne der mathematischen Wahrscheinlichkeitstheorie verhalten? Die Axiome (P1) - (P3) und ihre Folgerungen sind ja unsere statischen Gesetze für (die Koexistenz von) Glaubensgrade(n). Warum diese Gesetze und keine anderen? Es sind drei Argumente im Laufe der Zeit vorgetragen worden, die die Deutung subjektiver Glaubensgrade als Wahrscheinlichkeitsmaße zu einer Rationalitätsforderung machen. Es sind dies (1) Dutch Book-Theoreme, (2) Betrachtungen über komparative Wahrscheinlichkeiten und (3) Cox' Theorem. Das erste Argument ist das meist diskutierte; ich will es etwas ausführlicher behandeln; auf die andern gehe ich nur kursorisch ein.

(1) *Dutch Book-Theoreme*: Es ist mir nicht klar, wieso ausgerechnet die Holländer dafür herhalten müssen, als Leute bezeichnet zu werden, die andere übers Ohr hauen. Ein *Dutch Book* ist kein Buch, sondern ein System von Wetten, das ein Wetter mit einem Buchmacher eingeht, bei dem der Wetter verliert, egal wie die Ereignisse ausgehen, auf die gewettet wird. Und das Dutch Book-Theorem besagt, daß ein Wetter

---

[48] Um mal endlich Literatur zu nennen: Zur Geschichte des Wahrscheinlichkeitsbegriffs cf. Hacking (1975). Zur reinen Mathematik des Wahrscheinlichkeitsbegriffs s. Lipschutz (1976) (elementar) oder Bauer (1968) (sehr fortgeschritten). Für philosophische Einführungen sowohl in die elementare Mathematik wie in die Interpretationsproblematik s. Salmon (1967), Skyrms (1975). Fast die überzeugendste Möglichkeit, subjektive und objektive Deutungen des Wahrscheinlichkeitsbegriffs zusammenzuspannen, liefert meines Erachtens. Lewis (1986); cf. auch die Verbesserung in Lewis (1994).

für ein Dutch Book anfällig ist, wenn er seine Wahrscheinlichkeiten nicht im Einklang mit den Axiomen des Wahrscheinlichkeitskalküls, also mit den Gesetzen der statischen probabilistischen Glaubenstheorie ansetzt. Es gibt auch das umgekehrte Dutch Book-Theorem, das da lautet: wenn es möglich ist, ein Dutch Book zu konzipieren, dann kann dem Akteur eine Verletzung der Wahrscheinlichkeitsgesetze nachgewiesen werden.

Das ist zu erläutern. Hinter dem Theorem, das erstmals von Frank Ramsey in den 20er Jahren angedeutet wurde, steht die Beobachtung, daß wir unsere Handlungen dauernd von unseren subjektiven Wahrscheinlichkeiten leiten lassen; dauernd schätzen wir ab, wie wahrscheinlich sich wie gute und wie schlechte Folgen aus unseren Handlungen ergeben, so daß wir am Ende möglichst viel Glück und möglichst wenig Unglück in die Welt setzen. In aller Regel sind dabei Folgenbewertung und Wahrscheinlichkeitsabschätzung auf komplexe Weise verquickt.[49] Im Alltag sind Wetten zwar eher künstliche und marginale Situationen (unsere milliardenschweren Versicherungsgesellschaften sind freilich im Kern Wettkonzerne). Doch sind Wetten prototypisch für unsere permanente Befindlichkeit; sie bilden in der Tat besonders einfache Entscheidungssituationen, in denen die Folgenbewertung trivial ist und sich die Wahrscheinlichkeitsabschätzung in besonders reiner Weise ausdrückt. Deswegen sind sie theoretisch nützlich; und deswegen befasse ich mich im folgenden mit Wetten.[50]

Ist das Wetter morgen gut, oder wird es regnen? Wir können wetten: Ist es gut, zahlst du mir eine Mark, regnet es, zahle ich dir eine Mark. Aber nicht in allen Wetten müssen sich die Einsätze im Verhältnis 1:1 verhalten; es sind krumme Wetten möglich. Im Falle der eben betrachteten symmetrischen Wette beträgt der Gesamteinsatz zwei, der eigene Einsatz eine Mark. Sei allgemein a der Einsatz des Wetters und daher zugleich der Gewinn des Buchmachers, falls der Wetter verliert. Sei b der absolute Gewinn des Wetters im Erfolgsfall. Der Gesamteinsatz der Wette ist damit $a+b$. Das Verhältnis der Einsätze $a{:}b$ gibt den Buchmacherkurs an (engl. odds). Der Wettquotient für eine Wette ist $a/a+b$, also das Verhältnis vom Einsatz des Wetters zum Gesamteinsatz. Die Grundannahme ist nun, daß die subjektiven Wahrscheinlichkeiten sich in den Wettquotienten ausdrücken, die man einzugehen bereit ist.

---

[49] Die Theorie, die diese Verquickung en detail studiert, ist die Entscheidungstheorie, wozu von philosophischer Seite Jeffrey (1983) und von ökonomischer Seite Raiffa (1968) zu empfehlen sind.

[50] Der klassische Aufsatz, in dem die Idee von Graden rationalen Glaubens als rationalen Wettquotienten und die Methode von Dutch Book-Überlegungen zu ihrer Rechtfertigung erstmals auftauchen, ist: Ramsey (1929). Ein anderer Klassiker, mit dem ersten ausführlichen Beweis des Dutch Book-Theorems ist de Finetti (1937).

Allerdings bilden die Wettquotienten nicht den einzigen relevanten Faktor; in Abhängigkeit von der Risikobereitschaft der Wetter wird auch die Höhe der Gesamteinsätze eine erhebliche Rolle spielen. An Tagen, an denen sich ein Sommerhoch aufbaut, dürfte jeder von uns bereit sein, die oben erwähnte Wette über die Regenaussichten des folgenden Tages einzugehen. Doch was ist, wenn der Einsatz nicht bloß eine Mark, sondern eine Million Mark für jeden betragen soll? Dann bekommt man kalte Füße; Risikoscheu wird einen davon abhalten, diese Wette einzugehen. Um solche Probleme mit der Einstellung zum Risiko und mit der Nutzenbewertung von Geldbeträgen zu vermeiden, soll im weiteren angenommen werden, daß die Gesamteinsätze der Wetten sehr niedrig sind.

Als nächstes ist zu klären, was eine faire Wette ist. Nehmen wir an, es wird auf die Wahrheit von A gewettet. Sei S (= $a+b$) der Gesamteinsatz und r (= $a/a+b$) der Wettquotient. Dann ist die folgende Tabelle einschlägig:

	Einsatz	Auszahlung	Gewinn/Verlust
A	$r \cdot S$	S	$(1-r) \cdot S$
non-A	$r \cdot S$	0	$-r \cdot S$

Die Wette ist dann fair aus der subjektiven Sicht eines Wetters, wenn es ihm egal ist, welche Seite der Wette er annimmt. Das ist genau dann der Fall, wenn der mit der subjektiven Wahrscheinlichkeit $P(A)$ zu erwartende Gewinn im Fall der Wette auf A gleich dem mit $P(\text{non-}A)$ zu erwartenden Gewinn im Fall der Wette auf non-A ist, daß heißt, wenn $(1-r) \cdot S \cdot P(A) = r \cdot S \cdot P(\text{non-}A)$. Das wiederum gilt genau dann, wenn $r = P(A)$.[51] Eine Wette auf A ist also genau dann fair, wenn der Wettquotient gleich der subjektiven Wahrscheinlichkeit für A ist.

Bei Dutch Books passiert nun folgendes: dem Wetter wird ein ganzes System von Wetten angeboten, für die er die Wettquotienten festlegen darf. Der Buchmacher hingegen entscheidet über die Höhe der Gesamteinsätze und vor allem auch darüber, wer auf die Wahrheit von welcher Aussage wetten darf. Das zwingt den Wetter dazu, alle Wetten aus seiner Sicht fair zu gestalten. Doch muß der Wetter auch noch den Zusammenhang der verschiedenen Wetten des Systems beachten; er muß, wie es

---

[51] $(1-r) \cdot S \cdot P(A) = r \cdot S \cdot P(\text{non-}A)$ gdw. $(1-r) \cdot S \cdot P(A) = r \cdot S \cdot (1-P(A))$. Teilt man beide Seiten durch S, ergibt sich $(1-r) \cdot P(A) = r \cdot (1-P(A))$, d.h. $P(A) - r \cdot P(A) = r - r \cdot P(A)$. Addiert man auf beiden Seiten $r \cdot P(A)$, erhält man $P(A) = r$. Man beachte, daß beim ersten Schritt schon die Axiome der Wahrscheinlichkeitstheorie vorausgesetzt werden. Hält sich einer nicht daran, wird er Wetten auf der Basis fairer Wettquotienten nicht akzeptieren.

heißt, ein kohärentes System von Wetten eingehen. Ein Wettsystem heißt kohärent, wenn es kein Dutch Book ist, wenn sich also nicht in jedem Fall, wie immer die Wetten ausgehen, ein Verlust für den Wetter einstellt. Entsprechend heißen jemandes subjektive Wahrscheinlichkeiten genau dann kohärent, wenn es nicht möglich ist, ihn auf ihrer Basis in ein Dutch Book zu locken. Damit lassen sich die beiden Richtungen des Dutch Book-Theorems so ausdrücken: Die subjektiven Wahrscheinlichkeiten einer Person sind genau dann kohärent, wenn sie die Axiome (P1) - (P3) der mathematischen Wahrscheinlichkeit erfüllen.

Im folgenden Exkurs (den ein eiliger Leser überspringen kann) soll die eine Richtung des Dutch Book-Theorems bewiesen werden. Es lautet:

> Wenn ein Akteur seine subjektiven Wahrscheinlichkeiten so wählt, daß mindestens eines der Axiome (P1) - (P3) verletzt ist, so unterliegt er einem Dutch Book.

1. Teil: Angenommen, unser Akteur verletzt (P1), d.h. es ist $P(A) < 0$ für ein A. Da für faire Wetten Wahrscheinlichkeiten mit Wettquotienten identisch sind, gilt für die Wette auf A $r < 0$. Das bedeutet, daß genau einer der beiden Werte a (der Einsatz des Wetters) oder b (sein möglicher Gewinn) negativ ist. Ist a positiv und b negativ, so läßt der Buchmacher ihn auf A wetten, und dann passiert folgendes: Falls A wahr ist, bekommt er $S = a+b$ ausgezahlt. Sein „Gewinn" beträgt dann $-a+a+b < 0$. Falls A falsch ist, erhält er nichts, verliert also a. Ist a negativ und b positiv, so läßt der Buchmacher ihn auf non-A wetten, und die Sache geht genauso schlecht aus.

2. Teil: Angenommen, der Akteur verletzt (P2), d.h. die notwendige Proposition Ω hat nicht die Wahrscheinlichkeit 1. Zwei Unterfälle: (i) $P(\Omega) > 1$. Dann ist sein Wettquotient $r > 1$, also $a > a+b$, d.h. b ist negativ. Da Ω notwendigerweise wahr ist, gewinnt er die Wette auf jeden Fall und erleidet so den „Gewinn" b. (ii) $0 < P(\Omega) < 1$. Also ist $0 < r < 1$, damit $0 < a < a+b$. Dann läßt der Buchmacher ihn auf non-Ω, d.h. auf die unmögliche Proposition $\emptyset$ wetten; diese Wette und damit seinen Einsatz a verliert auf jeden Fall.

3. Teil: Das ist der komplizierte Teil des Beweises. Seien A und B miteinander unverträglich. Wir betrachten drei Wetten: Wette 1 auf A mit dem Gesamteinsatz S_1 und dem Wettquotienten r_1, Wette 2 auf B mit dem Gesamteinsatz S_2 und dem Wettquo-

tienten r_2 und Wette 3 auf $A \vee B$ (d.h. A oder B) mit dem Gesamteinsatz S_3 und dem Wettquotienten r_3. Nun sind folgende Fälle möglich:

(i) A und non-B tritt ein – Auszahlung: $(1–r_1) \cdot S_1 – r_2 \cdot S_2 + (1–r_3) \cdot S_3$,
weil der Wetter die erste und dritte Wette gewonnen hat, die zweite verloren.

(ii) non-A und B tritt ein – Auszahlung: $–r_1 \cdot S_1 + (1–r_2) \cdot S_2 + (1–r_3) \cdot S_3$,
weil der er die erste Wette verloren, die zweite und die dritte aber gewonnen hat.

(iii) non-A und non-B tritt ein – Auszahlung: $– r_1 \cdot S_1 – r_2 \cdot S_2 – r_3 \cdot S_3$,

weil der Wetter in diesem Fall alle drei Wetten verliert. Nehmen wir nun an, daß seine Wahrscheinlichkeiten und damit seine Wettquotienten das Axiom (P3) verletzen, daß also $r_1 + r_2 \neq r_3$. Und betrachten wir zunächst den Fall, daß $r_1 + r_2 > r_3$. Dann legt der Buchmacher die Einsätze mit $S_1 = S_2 = S > 0$ und $S_3 = –S$ fest. Daß damit ein Dutch Book entsteht, läßt sich leicht ausrechnen. Im Fall (i) etwa ist nach der Festsetzung des Buchmachers die Auszahlung

$$(1–r_1) \cdot S – r_2 \cdot S – S \cdot (1–r_3) = S \cdot (1 – r_1 – r_2 – 1 + r_3) = S \cdot (r_3 – r_1 – r_2) < 0.$$

Dasselbe ergibt sich, wie sich entsprechend nachprüfen läßt, im Fall (ii) und ebenso im Fall (iii), wenn alle drei Wetten verloren gehen. Zur Vervollständigung des Beweises ist noch der Fall zu behandeln, wo $r_1 + r_2 < r_3$ gilt. Nun setzt der Buchmacher die Gewinne mit $S_1 = S_2 = –S$ und $S_3 = S > 0$ fest; und dann läuft der Beweis läuft ganz analog. So weit der technische Exkurs.[52]

Das formale Dutch-Book-Theorem ist unanfechtbar. Das Dutch Book-Argument zur Rechtfertigung von (P1) - (P3) als Gesetze rationaler Glaubensgrade, geht aber darüber hinaus. Es geht von einer normativen Prämisse aus, nämlich daß es irrational ist, sich einem Dutch Book aussetzen. Und aus dieser normativen Prämisse erhalten wir mit all den Annahmen über die Wettsituation und dem formalen Theorem

[52] Der Beweis ist eine ausgeführte Form der Skizze bei Earman (1992). Eine gute Darstellung von Wettsystemen und dem Dutch Book-Theorem findet man in Skyrms (1975). Andere ausführliche Darstellungen geben Resnik (1987) oder die fast schon klassischen Aufsätze von Shimony (1955), Kemeny (1955), auf welch letzteren Aufsatz für einen Beweis des umgekehrten Dutch Book-Theorems verwiesen sei.

schließlich die normative Konklusion, daß die Glaubensgrade rationalerweise (P1) - (P3) erfüllen, sich also wie mathematische Wahrscheinlichkeiten verhalten.[53]

(2) *Komparative Wahrscheinlichkeiten*: Wem das Dutch-Book-Argument nicht gut genug ist, kann auch die folgende Gedankenlinie einschlagen. Dazu bedarf es erst einmal einiger Definitionen. Wir schreiben „$A \ll B$" für „B ist sicherer als A" und „$A \underline{\ll} B$" für „B ist mindestens so sicher wie A". Diese beiden Relationen hängen auf einfache Weise zusammen: $A \ll B$ genau dann, wenn $A \underline{\ll} B$ und nicht $B \underline{\ll} A$. Mit diesen komparativen Relationen lassen sich Glaubensgrade für Propositionen wenigstens vergleichen (was etwas Schwächeres ist als sie zu spezifizeren). Folgende komparative Eigenschaften von Glaubensgraden erscheinen dann zumindest plausibel:

(K1) $\emptyset \ll \Omega$,
(K2) wenn $A \underline{\ll} B$ und $B \underline{\ll} C$, dann $A \underline{\ll} C$ (Transitivität),
(K3) $A \underline{\ll} B$ oder $B \underline{\gg} A$ (Vollständigkeit),
(K4) wenn A mit B und mit C unverträglich ist, so gilt: $B \underline{\ll} C$ gdw. (A oder B) $\underline{\ll}$ (A oder C).

Damit läßt sich das folgende (in der Tat sehr einfache) Theorem beweisen:

> (K1) - (K4) sind notwendige Bedingungen dafür, daß es ein Wahrscheinlichkeitsmaß P gibt, so daß $P(A) \leq P(B)$ gdw. $A \underline{\ll} B$.

Deswegen heißen (K1) - (K4) auch Axiome der komparativen Wahrscheinlichkeit. Da sie aber nur notwendige Bedingungen für die genannte Aussage sind, liefert ihre Plausibilität noch keine Rechtfertigung der quantitativen Wahrscheinlichkeitsaxiome. Nun gibt es aber auch Versionen dieses Theorems, in denen hinreichende Bedingungen für die Existenz eines mit den komparativen Relationen übereinstimmenden Wahrscheinlichkeitsmaßes angegeben werden; grob gesagt, sind damit die Axiome (P1) - (P3) aus diesen Bedingungen ableitet.[54] Wiederum ist dieser mathematische Sachverhalt allein noch keine normative Rechtfertigung dieser Axiome. Vielmehr muß man sich diese hinreichenden Bedingungen genau daraufhin anschauen, ob sie

[53] Es ist keineswegs so, daß die Dutch-Book-Argumentation (im Unterschied zum Theorem) unumstritten wäre; die Einwände richten sich gegen die Rationalitätsprämisse (die mir aber einleuchtend erscheint) oder gegen bestimmte Annahmen über die Akzeptanz von Wettsystemen (die eher angreifbar sind). Vgl. etwa Maher (1993). Kritische Anmerkungen gibt es auch bei Earman (1992).

[54] Die Sache wird dann aber mathematisch recht anspruchsvoll. Für eine genaue formale Darstellung vgl. Krantz, Luce, Suppes und Tversky (1971) oder Stegmüller (1973).

ebenso plausibel und rational rechtfertigbar sind wie (K1) - (K4); erst dann läßt sich die Rechtfertigung mithilfe des Theorems auf (P1) - (P3) übertragen.

(3) *Cox' Theorem*: Dieses liefert das letzte Argument zur Verteidigung von Wahrscheinlichkeitsmaßen als Repräsentationen subjektiver Überzeugungsgrade, das ich hier nennen möchte. Unter Philosophen blieb es weitgehend unbekannt; das Resultat wird meist in der Künstlichen Intelligenz-Forschung erwähnt. (KI-Forscher sind ja ebenfalls an erkenntnistheoretischen Fragen interessiert. Sie haben die Aufgabe, geeignete Modellierungen epistemischer oder doxastischer Systeme zu finden, die dann von Computern benutzt werden können, sehen in der Wahrscheinlichkeitstheorie dafür ein geeignetes Mittel und stehen daher vor demselben Rechtfertigungsproblem.) Die Grundidee hinter Cox' Theorem ist recht verschieden von den bisherigen Argumentationen; deshalb möchte ich sie wenigstens kurz erwähnen. Zunächst sei angenommen, daß ein doxastischer Zustand jedenfalls durch irgendeine Zuordnung von numerischen Glaubensgraden zu Propositionen repräsentiert ist. Sei Q eine solche Zuordnung, d.h. eine Funktion, die jeder Teilmenge von Ω eine reelle Zahl zuweist. Damit die Werte von Q als Glaubensgrade verstanden werden können, scheinen wenigstens die folgenden zwei Minimalbedingungen gelten zu müssen:

(C1) es gibt eine nicht steigende Funktion h, so daß $Q(\text{non-}A) = h(Q(A))$,

(C2) es gibt eine in beiden Argumenten nicht fallende Funktion f, so daß $Q(A \text{ und } B) = f(Q(A), Q(B \mid A))$.

(C1) sagt also, daß der Glaubensgrad von non-A von dem von A funktional abhängt und daß, wenn letzterer steigt, ersterer nicht auch steigen kann. (C2) sagt, daß (A und B) jedenfalls nicht unsicherer wird, wenn A sicherer oder B unter der Bedingung A sicherer wird, und ferner, daß der Glaubensgrad von (A und B) funktional von den beiden anderen Glaubensgraden abhängt. Diese Bedingungen sehen sehr schwach aus und sind zweifelsohne plausibel. Wahrscheinlichkeitsmaße erfüllen ersichtlich (C1) und (C2); bei ihnen ist $h(x) = 1 - x$ und $f(x,y) = x \cdot y$. Fügt man dem noch einige weitere Bedingungen hinzu – die eher technischer Natur zu sein scheinen, die man sich aber doch genau anschauen müßte –, so kann man Cox' Theorem formulieren: Gelten (C1), (C2) und diese weiteren Bedingungen, dann muß Q ein Wahrscheinlichkeitsmaß sein.

Der Witz von Cox' Theorem gleicht also dem des Satzes über die komparative Wahrscheinlichkeiten: aus schwachen Annahmen über Überzeugungsgrade, die nach Möglichkeit jeder akzeptieren kann, ergibt sich zwingend eine Repräsentation doxastischer Zustände als Wahrscheinlichkeitsmaße. Doch wieder hängt die Überzeugungskraft des Arguments an der Überzeugungskraft dieser Annahmen.[55]

So viel zur Begründung der Wahrscheinlichkeitsaxiome (P1) - (P3) in ihrer subjektivistischen Interpretation. Zwei weitere Fragen verdienen in diesem Zusammenhang noch unsere Aufmerksamkeit. Die eine Frage ist, wie sich die probabilistische zur qualitativen Beschreibungsweise verhält, die wir zuvor betrachtet haben. Die Antwort scheint nahezuliegen. Die probabilistische Beschreibungsweise ist ja offensichtlich viel feiner differenziert als die qualitative, und daher könnte man denken, daß der qualitative Glauben nichts anderes bedeutet als (sehr) hohe Wahrscheinlichkeit. Formaliter läuft das auf die folgende Definition oder Reduktion hinaus:

$$G(A) \text{ gdw. } P(A) \geq 1\text{-}\varepsilon, \text{ für geeignet kleines } \varepsilon.$$

Doch ist das inadäquat, wie man aus dem sogenannten Lotterieparadox ersehen kann: Wie immer ε festgesetzt wird, es kann passieren, daß diese Reduktion zu einem inkonsistenten Glaubenssystem führt. Dazu betrachten wir eine Lotterie mit n Losen, von denen exakt eines gewinnt; dabei sei $1/n < \varepsilon$. Dann gilt für meine Wahrscheinlichkeiten: P(Los 1 gewinnt nicht) = $1 - 1/n$. Also glaube ich, daß Los 1 nicht gewinnt. In gleicher Weise gilt für jedes Los k ($k = 1,...,n$): ich glaube, daß Los k nicht gewinnt. Doch außerdem glaube ich natürlich, daß eines der n Lose jedenfalls gewinnt. Damit habe ich aber inkonsistente Überzeugungen, im Widerspruch zu den Axiomen der epistemischen Logik.

Man kann die Untauglichkeit der Reduktionsidee auch anders nachweisen. Aus $G(A)$ und $G(B)$ folgt gemäß (G2), daß $G(A \wedge B)$. Wäre die obige Reduktion korrekt, so müßte gelten: aus $P(A) \geq 1 - \varepsilon$ und $P(B) \geq 1 - \varepsilon$ folgt $P(A \wedge B) \geq 1 - \varepsilon$. Das folgt aber überhaupt nicht. Wenn etwa A und B probabilistisch unabhängig sind, so gilt $P(A \wedge B) = P(A) \cdot P(B)$; und dieses Produkt kann leicht unter $1 - \varepsilon$ sinken, auch wenn beide Faktoren noch darüber liegen. Qualitativer Glaube und Wahrscheinlichkeiten verhalten sich in Bezug auf Konjunktionen ganz verschieden.

[55] Cox' eigene Darlegung des Sachverhalts findet sich am besten in Cox (1961). Die im Text angedeutete Version des Theorems findet man bei Fine (1973). Zu einer Variante mit etwas anderen Axiomen samt einer kritischen Diskussion cf. Earman (1992).

So geht die Reduktion also nicht. Angesichts dieser Probleme, könnte man denken, daß der Fehler daher rührt, daß man qualitativen Glauben nur mit hoher Wahrscheinlichkeit gleichgesetzt hat (wo $\varepsilon > 0$); vielleicht sollte man ihn besser mit maximaler Wahrscheinlichkeit gleichsetzen:

$$G(A) \text{ gdw. } P(A) = 1.$$

Damit verschwinden die genannten Probleme. Doch paßt diese Festsetzung schlecht in die dynamische Glaubenstheorie. Wie wir bald sehen werden, gilt dort, daß eine Proposition, die einmal die Wahrscheinlichkeit 1 erhält, diese niemals wieder verliert. Für Wahrscheinlichkeiten mag dies in Ordnung sein. Für qualitative Überzeugungen gilt dies aber sicherlich nicht; die meisten unserer Überzeugungen sind mehr oder weniger leicht revidierbar. So kann man momentan nur eins sagen: das Verhältnis von qualitativer und probabilistischer Beschreibung der statischen Gesetze für rationale Überzeugungen ist bisher nicht befriedigend geklärt. Offenbar haben beide Konzeptionen ihre Berechtigung; und schon deswegen müssen beide parallel diskutiert werden.

Die andere Frage ist, ob es denn über die Axiome (P1) - (P3) hinaus noch weitere plausible Rationalitätspostulate an Glaubensgrade gibt. Dazu gibt es in der Tat eine umfängliche Diskussion, die ich hier nur andeuten kann. Ein ziemlich einleuchtendes Postulat ist das von Carnap vorgeschlagene Regularitätsaxiom, welches das Theorem (P4) umkehrt:[56]

$$P(A) = 0 \text{ nur dann, wenn } A = \emptyset.$$

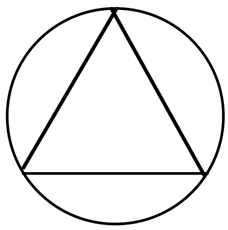
Es ist sicher auch nicht vernünftig, sich auf Wettsysteme einzulassen, bei denen keine Möglichkeit für einen Gewinn, wohl aber eine für einen Verlust besteht. Nennen wir ein Wettsystem strikt kohärent, wenn es möglich ist, daß der Wetter auch gewinnt. Unter strikter Kohärenz muß auch das Regularitätsaxiom gelten. Denn wer bereit ist, auf eine kontingente Proposition zu wetten derart, daß er bei ihrer Wahrheit nichts bekommt, ist nicht strikt kohärent.

Weitergehende Rationalitätspostulate werden unter der Bezeichnung „Symmetrieaxiom“ oder (etwas ironisierend) „Prinzip des unzureichenden Grundes“ abge-

[56] Cf. z.B. Carnap und Stegmüller (1959).

handelt. Angenommen, wir haben einen Würfel, von dem wir gar nichts wissen. Vielleicht ist er fair, vielleicht gezinkt; aber wenn er gezinkt ist, so haben wir keine Ahnung, in welche Richtung er gezinkt sein könnte. Das Axiom sagt nun, daß wir – solange wir keinen Grund haben, über einen der möglichen Fälle anders zu denken als über die andern – alle Fälle gleich behandeln, d.h. ihnen die gleiche Wahrscheinlichkeit zuweisen sollten.

Doch verwickelt man sich mit diesem Symmetrieaxiom leicht in Schwierigkeiten. Wie, sei mit einer berühmten Kopfnuß für Mathematiker illustriert: Bertrands Paradox. Dazu betrachte man die nachfolgende Figur:



Auf diesen Kreis lasse man zufällig einen Stock so fallen, daß er eine Sehne durch den Kreis legt. Frage: Wie groß ist die Wahrscheinlichkeit, daß diese Sehne länger ist als die Seiten des einbeschriebenen gleichseitigen Dreiecks? Zur Beantwortung scheint man sich auf drei verschiedene Weisen auf das Symmetrieaxiom berufen zu können, die aber alle verschiedene Resultate zeitigen:

(a) Alle Richtungen, in denen der Stock auf dem Kreis zu liegen kommt, sind gleichberechtigt. Nehmen wir also an, der Stock liege waagrecht, und betrachten die senkrechte Linie durch den Mittelpunkt des Kreises. Der Mittelpunkt der vom Stock markierten Sehne liegt auf dieser Linie. Wiederum scheint jeder Punkt auf dieser Linie mit gleichem Recht dieser Mittelpunkt sein zu können. Liegt nun dieser Sehnenmittelpunkt näher zum Kreismittelpunkt als die Dreiecksseite, so ist die Sehne länger als die Dreiecksseite, ansonsten kürzer. Der senkrechte Radius wird aber von der Dreiecksseite gerade halbiert. Also ist dem Symmetrieaxiom zufolge die Wahrscheinlichkeit, daß der Stock eine längere Sehne ausschneidet, gleich 1/2.

(b) Alle Punkte, in denen der Stock den Kreis schneidet, sind gleichberechtigt. Sei also die Spitze des Dreiecks einer der beiden Schnittpunkte des Stocks mit dem Kreis. Legen wir nun in diesem Punkt eine Tangente an den Kreis. Die ausgeschnittene Sehne nimmt dann von diesem Punkt aus eine beliebige Richtung innerhalb der 180°

unter der Tangente. Die Sehne ist länger als die Dreiecksseite, wenn der Stock innerhalb des Dreieckswinkels von 60° liegt; sonst ist sie kürzer. Da alle Richtungen gleichwahrscheinlich sind, diktiert das Symmetrieaxiom nun 1/3 als Antwort.

(c) Die letzte Überlegung führt zu der Antwort 1/4. Die vom Stock ausgeschnittene Sehne hat irgendwo im Kreis ihren Mittelpunkt. Denken wir uns nun den dem Dreieck einbeschriebenen Kreis hinzu. Liegt der Sehnenmittelpunkt in diesem kleinen Kreis, so ist die Sehne länger als die Dreiecksseiten; ansonsten kürzer. Der kleine Kreis hat den halben Durchmesser des großen und mithin nur ein Viertel der Fläche. Dem Symmetrieaxiom zufolge sind alle Punkte im großen Kreis, als Sehnenmittelpunkt gleichwahrscheinlich; daher die dritte Antwort.

Hier geht etwas schief; offenbar machen wir hier vom Symmetrieaxiom zu freizügig Gebrauch. Welches nun die richtige Antwort ist, ist eine vertrackte wahrscheinlichkeitstheoretische Frage. Hier soll diese Geschichte nur Schwierigkeiten mit dem Symmetrieaxiom aufzeigen. So mancher, etwa der Miterfinder der Dutch Book-Argumentation, Bruno de Finetti, hat aus solchen Geschichten die Konsequenz gezogen, daß die Wahrscheinlichkeitsaxiome (P1) - (P3) die einzigen Rationalitätspostulate für subjektive Wahrscheinlichkeiten sind.[57]

Wie schon erwähnt, nimmt die Wahrscheinlichkeitstheorie unter allen Theorien der Glaubensgrade einen absolut vorherrschenden Platz ein. Das hat historische Gründe, ist aber angesichts der gelieferten Begründungen vielleicht auch sachlich berechtigt. Doch gibt es, wie auch schon erwähnt, mittlerweile etliche Alternativen, und auf zwei davon möchte ich wenigstens hinweisen. Da gibt es zum einen die Dempster-Shafer-Theorie der „belief functions“. Ihre Idee besteht in der folgenden Generalisierung: wir nehmen nicht mehr an, daß die Glaubensgrade in eindeutigen bestimmten Wahrscheinlichkeiten bestehen, sondern bloß, daß sie in mehr oder weniger großen Intervallen von Wahrscheinlichkeiten bestehen; die klassische Wahrscheinlichkeitstheorie ist dann ein Spezialfall dieses allgemeineren Vorgehens.[58] Bi-

---

[57] Eine Diskussion von Bertrands Paradox mit weiteren Literaturangaben und Hinweisen auf die Geschichte des Problems enthält Marinoff (1994). Marinoff ist der Ansicht, daß das Problem durch eine unzureichend präzise Formulierung entsteht (der wir uns in unserer Darstellung auch schuldig machen) und es für alle Antworten mindestens eine Präzisierung der Problemstellung gibt, auf die sie die richtige Antwort sind. Cf. zu Bertrands Paradox und dem Symmetrieaxiom auch van Fraassen (1989).

[58] Ausgearbeitet wurde diese Theorie durch Shafer (1976). Darin werden auch einige von Philosophen und Ökonomen (Rescher, Shackle, Cohen) entwickelte Alternativkonzeptionen diskutiert. In eher historisch gerichteten Arbeiten verweist Shafer auf die interessante Tatsache, daß solche Alternativen eigentlich schon während der gesamten letzten 300 Jahre auf dem Markt waren. Cf. Shafer (1978).

zarrer als diese Modelle sind die Ideen der von L.A. Zadeh entwickelten „fuzzy logic“, wo man mit graduell abgestuften Wahrheitswerten arbeitet, die sich aber nicht wie Wahrscheinlichkeiten verhalten und vielleicht auch als Glaubensgrade interpretieren lassen.[59]

Diese Alternativen werden in der KI-Forschung besonders intensiv verfolgt. Das hat Gründe. In der KI will man ja so etwas wie künstliche epistemische Systeme bauen, z.B. ärztliche Diagnosesysteme. Verwendet man dort Wahrscheinlichkeitsmaße zur Repräsentation doxastischer Zustände, so ist man sehr schnell dabei, die Kapazität der Rechner nachhaltig zu überfordern. Denn im Prinzip ist ein Wahrscheinlichkeitsmaß erst dann spezifiziert, wenn man jeder Proposition eine Wahrscheinlichkeit zugeordnet hat. Gemäß (P9) läßt sich das darauf zurückführen, daß man allen einzelnen Fällen Wahrscheinlichkeiten zuordnet. Und diese müssen sich zu 1 addieren; das ist die einzige Beschränkung, der sie unterliegen. Umfaßt der Möglichkeitsraum n Fälle, so gibt es also n–1 Freiheitsgrade in der Spezifikation eines Wahrscheinlichkeitsmaßes. Doch wachsen die Möglichkeitsräume exponentiell. Wenn man es in einer Anwendung etwa mit k binären, logisch unabhängigen Variablen zu tun hat, so gibt es 2^k Möglichkeiten, wie sich diese Variablen realisieren können. Für die Verteilung von Wahrscheinlichkeiten darüber gibt es also $2^k - 1$ Freiheitsgrade. Und mit jeder neu hinzukommenden binären Variable verdoppelt sich die Anzahl der Möglichkeiten. Ein Diagnosesystem etwa, das 100 Variablen berücksichtigt – und für reale Anwendungen sind das eher wenig –, braucht mithin $2^{100} - 1$ unabhängige Wahrscheinlichkeiten. Damit ist jeder Computer hoffnungslos überfordert. Also sucht man nach komputationell handhabbaren Alternativen. Und so ist dort nach anfänglichem Herumtappen eine blühende Forschung entstanden, aus denen sich auch die Erkenntnistheoretiker einiges mit Gewinn abschauen können. Ich will dies im nächsten Kapitel anhand der Theorie der Bayes'schen Netze vorführen, welche die wichtigste Methode bildet, das gerade geschilderte komputationelle Problem im Rahmen der Wahrscheinlichkeitstheorie in den Griff zu bekommen.

59 Vgl. etwa Dubois und Prade (1988).

5. Dynamische Glaubenstheorie

Die dynamische Glaubenstheorie ist der philosophisch spannendere und bedeutsamere Teil, den man freilich nur auf der Grundlage einer gegebenen statischen Theorie, wie wir sie gerade diskutiert haben, ausarbeiten kann. Im dynamischen Teil geht es um die Frage, wie sich Überzeugungen oder doxastische Zustände im allgemeinen im Laufe der Zeit verändern. Um die philosophische Relevanz dessen zu begreifen, muß man verstehen, daß eine Beantwortung dieser Frage der einzige Weg zu einer konstruktiven Behandlung des klassischen Induktionsproblems ist. Dies sei zunächst erläutert:

Dem Induktionsproblem sind wir schon bei der Erläuterung der Humeschen Induktionsskepsis im Kapitel 2 begegnet. Es besteht in seiner allgemeinsten Formulierung in der Frage, wie und mit welcher Rechtfertigung wir aus dem Input, den wir bekommen, unser weit über den Input hinausgreifendes Weltbild – sei es in Bezug auf Allgemeines, auf die Zukunft, auf nicht überbrückte Distanzen, etc. – gewinnen. Nennen wir eine solche Methode ein induktives Schema. Hume verlangte hier eine deduktive Rechtfertigung, sah ein, daß sie nicht zu haben ist, und fand dann keine schwächere Begründung. Danach suchen wir auch jetzt nicht; um Begründungen geht es im nächsten Kapitel. Im Moment ist nur festzustellen, daß das Induktionsproblem keine dynamische Frage zu sein scheint. Sie ist aber darin versteckt:

Die dynamische Frage, wie sich unsere doxastischen Zustände ändern, läßt vorderhand viele Antworten zu. Man erfährt vieles, durch die eigenen Sinne und vom Hörensagen; man denkt nach und macht sich Dinge klar; man vergißt laufend etwas; man betrinkt sich, mit beträchtlichen Folgen für den kognitiven Zustand; manchmal beruhen Überzeugungen auf purem Wunschdenken; und so weiter, da gibt es viele Ursachen. Zu den meisten hat der Philosoph als solcher nichts zu sagen. Aber manche lassen sich einer Rationalitätsbeurteilung unterwerfen; und diese sind Gegenstand philosophischen Interesses. Dazu gehören insbesondere die doxastischen Änderungen aufgrund neuer Wahrnehmungen, Erfahrungen oder Informationen; nur mit diesen will ich mich im weiteren beschäftigen.[60] Was wir also als Antwort auf die dynami-

[60] Nicht weil es keine anderen gäbe – die Ergebnisse von Nachdenken, von begrifflichen Erweiterungen u.ä. lassen sich durchaus als rational beurteilen, und vielleicht gibt es so etwas wie rationales Vergessen – und nicht weil diese unwichtig wären; es ist nur so, daß es dazu nach meinem Überblick keine guten theoretischen Vorstellungen gibt.

sche Frage suchen, ist ein irgendwie als rational ausweisbares Änderungsschema, wie ich es nennen will, d.h eine Funktion, die einem beliebig gegebenen doxastischen Ausgangszustand und einer beliebig gegebenen hinzutretenden Information den daraus resultieren doxastischen Endzustand zuordnet.

Der springende Punkt ist nun, daß ein induktives Schema und ein Änderungsschema im wesentlichen auf dasselbe hinauslaufen. Ein induktives Schema projiziert aus den bis zu einem gegebenen Zeitpunkt vorliegenden Daten den doxastischen Ausgangszustand und aus diesen und den bis zu einem späteren Zeitpunkt hinzutretenden Daten den zugehörigen Endzustand; auf diese Weise erzeugt es ein Änderungsschema. Umgekehrt erzeugt ein Änderungsschema zusammen mit einem vor aller Erfahrung gegebenen Apriori-Zustand ein induktives Schema, da es dann jeder Datenmenge in einem oder in vielen Schritten den zugehörigen Endzustand zuordnet. Zwei Unterschiede gibt es nur: Zum einen kommt ein Änderungsschema nicht ohne den Apriori-Zustand als allerersten Startpunkt aus – womit man das zusätzliche Problem hat, diesen Apriori-Zustand zu charakterisieren –, während der Induktionstheoretiker sich auf der Suche nach einem induktiven Schema von diesem Problem unabhängig wähnte. Zum andern hoffte der Induktionstheoretiker traditionellerweise, ein einziges induktives Schema irgendwie als rational auszeichnen zu können, während die dynamische Theorie ohne diesen Eindeutigkeitsanspruch nach (gewiß möglichst starken) Rationalitätsbedingungen sucht. Aufgrund dieser Überlegung sehe ich in der Entwicklung einer dynamischen Glaubenstheorie auch und gerade eine aufgeklärte Behandlung des Induktionsproblems.

Die relativ ausgefeilteste dynamische Theorie gibt es, wie nicht überrascht, im Rahmen der probabilistischen Beschreibungsweise doxastischer Zustände. Wenden wir uns also zuerst diesem Rahmen zu.

Die elementarste Regel zur Änderung doxastischer Zustände ist dort die der *einfachen Konditionalisierung*. Sie sagt folgendes: Zum Zeitpunkt t hat eine Person das Wahrscheinlichkeitsmaß P, das seine Überzeugungsgrade repräsentiert. Zu einem späteren Zeitpunkt t' hat sie ein anderes Wahrscheinlichkeitsmaß P'. Zwischen t und t' hat sie die Erfahrung E gemacht oder die Information E erhalten (es geht, wie gesagt, nur um auf diese Weise induzierte Änderungen). Wie hängen P und P' zusammen? Gemäß der Regel der einfachen Konditionalisierung gilt für alle Propositionen A:

$$P'(A) = P(A|E).$$

Die neue oder, wie sie auch heißt, Aposteriori-Wahrscheinlichkeit für A ist danach einfach die durch die Erfahrung E bedingte alte oder Apriori-Wahrscheinlichkeit von A. Vorausgesetzt ist dabei $P(E) > 0$; ansonsten ist die Regel nicht definiert.

Diese Regel klingt so selbstverständlich, daß man erst spät auf die Idee kam, sie explizit zu formulieren. Daß sie aber sehr alt ist, sieht man an dem in der Statistik schon lange benutzten, sogenannten Theorem von Bayes: Sei die Menge $\{A_1,..., A_n\}$ eine Zerlegung oder Partition von Ω; das bedeutet, daß die Adjunktion aller A_i gleich Ω ist, und alle A_i einander wechselseitig ausschließen, daß also genau eine dieser Propositionen wahr ist. Dann gilt zunächst die Formel von der totalen Wahrscheinlichkeit:

$$P(E) = \sum_{i=1}^{n} P(E|A_i) \cdot P(A_i)$$

Schnell der Beweis: Da E mit $(E \wedge A_1) \vee ... \vee (E \wedge A_n)$ logisch äquivalent ist und alle Disjunktionsglieder einander wechselseitig ausschließen, gilt:

$$P(E) = P(E \wedge A_1) + ... + P(E \wedge A_n) = \sum_{i=1}^{n} P(E \wedge A_i) = \sum_{i=1}^{n} P(E|A_i) \cdot P(A_i).$$

Mit diesem Gesetz kann man bereits einiges anfangen. Man stelle sich vor, mit einer Münze konfrontiert zu werden, von der man nicht weiß, ob sie normal ist oder nur Kopf erlaubt. Sei E das Ergebnis „Kopf", A_1 die Hypothese, daß es eine normale Münze ist, und A_2 die Hypothese, daß sie auf beiden Seiten einen Kopf trägt. Dann kenne ich die bedingten Wahrscheinlichkeiten: $P(E|A_1) = 1/2$, $P(E|A_2) = 1$. Wenn ich nun subjektive Wahrscheinlichkeiten für das Vorliegen der beiden Hypothesen habe, kann ich daraus mit der obigen Formel $P(E)$ berechnen.

Ein Schritt weiter führt zu Bayes' Theorem. Sei wieder eine Zerlegung $\{A_1,...,A_n\}$ von Ω gegeben; in der üblichen Interpretation ist das eine Menge statistischer Hypothesen, von denen genau eine zutrifft. Nun wollen wir wissen, wie sich $P(A_j)$ im Lichte neuer Daten E verändert; gesucht wird also $P(A_j|E)$. Definitionsgemäß ist das einfach $P(E \wedge A_j) / P(E)$. Wenden wir nun auf den Nenner dieses Ausdruck die Formel von der totalen Wahrscheinlichkeit und auf den Zähler das Multiplikationsprinzip an, so erhält man Bayes' Theorem:

$$P(A_j|E) = \frac{P(E|A_j)\cdot P(A_j)}{\sum_{i=1}^{n} P(E|A_i)\cdot P(A_i)}.$$

Der Ausdruck $P(E|A_j)$, also die Wahrscheinlichkeit der vorgefundenen Daten aufgrund der jeweiligen Hypothese wird oft (auch in der deutschsprachigen Literatur) als die Likelihood der Hypothese bezeichnet. Bayes' Theorem berechnet also die Aposteriori-Wahrscheinlichkeit einer Hypothese aus den Apriori-Wahrscheinlichkei-ten aller Hypothesen und ihren Likelihoods; am einfachsten kann man es in der Aussage zusammenfassen, daß diese Aposteriori-Wahrscheinlichkeit proportional zur Apriori-Wahrscheinlichkeit und zur Likelihood sei. Für die sogenannten Bayesianer unter den Statistikern ist dieses Theorem deswegen zum Grundgesetz aller Statistik geworden, weil die rechte Seite der Gleichung lauter Größen enthält, die sie als gegeben betrachten können.[61] In diesem Theorem ist nicht von Zeitpunkten die Rede; es formuliert als keine explizite Dynamik. Doch war die Interpretation immer eine dynamische, wonach es schlicht auf die einfache Konditionalisierung hinausläuft.

Die einfache Konditionalisierung weist aber einige Nachteile auf. Der erste besteht darin, daß sie nicht anwendbar ist, wenn $P(E) = 0$. Dem kann man abhelfen, indem man zusätzlich das Regularitätsaxiom fordert. Eine andere Möglichkeit besteht darin, von vornherein mit bedingten Wahrscheinlichkeiten zu arbeiten, aus denen sich absolute Wahrscheinlichkeiten ableiten lassen. Daraus resultiert eine etwas andere Begrifflichkeit, in der man auch Bedingungen zulassen kann, die die absolute Wahrscheinlichkeit 0 haben. Der erste, der so etwas entwickelte, war Karl Popper, weshalb solche Wahrscheinlichkeitsmaße auch Popper-Maße heißen.[62]

Ein anderer, gewichtigerer Einwand lautet, daß die neue Erfahrung die Wahrscheinlichkeit 1 erhält: $P'(E) = P(E|E) = 1$. Aber wenn etwas einmal die Wahrscheinlichkeit 1 erhalten hat, kann es gemäß der einfachen Konditionalisierung diesen Wert nicht mehr verlieren. Demgegenüber scheint man nicht leugnen zu können, daß es auch unsichere Information gibt, die aber auch die Wahrscheinlichkeiten verändert. Das klassische Beispiel stammt von Richard Jeffrey: Ich versuche, bei Kerzenlicht

---

[61] Man muß hier aufpassen und zwischen *Bayes' Theorem* und *Bayes' Prinzip* unterscheiden. Das Theorem ist, wie gerade bewiesen, ein Theorem der Wahrscheinlichkeitstheorie; das Prinzip besagt, daß man die Handlung wählen soll, die den erwarteten Nutzen maximiert (cf. dazu Kapitel 8), und ist insofern ein Postulat der praktischen Rationalität, aber kein Theorem. Beides geht auf Bayes (1764) zurück. Die oft verwandte Bezeichnung „Bayesianer“ ist etwas unscharf. In der Statistik sind es die, die sich auf Bayes' Theorem berufen, in der Philosophie ganz allgemein die, die meinen, für die Beschreibung doxastischer Zustände und die Erkenntnistheorie im allgemeinen reichten probabilistische Mittel hin.

[62] Die erste Formulierung dieser Idee findet sich in Popper (1934), Neuer Anhang IV.9. Parallel wurde sie entwickelt von Renyi (1955). Cf. auch Spohn (1986).

die Farbe von Gegenständen zu ermitteln. Dieses Tuch ist grün oder blau, das weiß ich. Doch ist bei Kerzenschein nicht definitiv zu entscheiden, ob es nun grün oder blau ist; immerhin sieht es eher so aus, als wäre es blau. Hier habe ich eine unsichere Information erhalten. Wenn nun von meiner Einschätzung der Farbe andere Meinungen von mir abhängen, wie haben sich dann diese aufgrund dieser Information geändert? Jeffrey hat einen Vorschlag, wie man dieses Problem lösen kann; er wird als *allgemeine Konditionalisierung* oder *Jeffrey-Konditionalisierung* bezeichnet:[63]

Das allgemeine Szenario ist nun folgendes: Zu t hat eine Person das Wahrscheinlichkeitsmaß P. Im Zeitraum von t zu t' macht sie Erfahrungen, die nur die Propositionen in der Zerlegung $\{E_1,...,E_n\}$ betreffen. Die Erfahrungen lassen sie aber durchaus darüber unsicher, welches dieser E_i eingetreten ist. Zu t' hat er nach diesen Erfahrungen ein neues Wahrscheinlichkeitsmaß P'. Wie geht P' aus P hervor? Jeffrey geht hier zunächst von der Formel von der totalen Wahrscheinlichkeit in Bezug auf die alten Wahrscheinlichkeiten aus, die im vorliegenden Fall für jede Proposition A

$$P(A) = \sum_{i=1}^{n} P(A|E_i) \cdot P(E_i)$$

liefert. Sodann nimmt er an, daß sich die alten, durch die Propositionen E_i dieser Zerlegung bedingten Wahrscheinlichkeiten durch diese Erfahrungen nicht geändert haben; lediglich die Einschätzung dieser Propositionen hat sich geändert. Dies ergibt unmittelbar die folgende Regel:

$$P'(A) = \sum_{i=1}^{n} P(A|E_i) \cdot P'(E_i) \qquad \text{(Jeffrey-Konditionalisierung)}$$

Die einfache Konditionalisierung ist ein Spezialfall dieser Regel; wenn nämlich eines der E_i sicher ist, etwa E_1, dann gilt natürlich $P(E_1) = 1$ und $P(E_i) = 0$ für alle i = 2,3...n, womit sich Jeffreys Formel auf die der einfachen Konditionalisierung reduziert.

So weit habe ich die zwei hauptsächlich diskutierten probabilistischen Änderungsregeln nur erläutert. Die normative Frage, warum man sich gerade an diese und keine anderen Regeln halten sollte, warum gerade diese Form von Änderung rational ist, ist damit aber nicht beantwortet. Dafür gibt es meinem Überblick zufolge drei unterscheidbare Argumente.

---

[63] Cf. Richard Jeffrey (1983).

Die erste und wichtigste Antwort ist wiederum die Dutch-Book-Argumentation, die nun aber mit Wettsystemen operiert, die auf mehrere Zeitpunkte verteilt sind. Der Buchmacher bietet dem Wetter zu t und zu t' Systeme von Wetten an, die so beschaffen sind, daß genau dann, wenn der Wetter seine Überzeugungen im Einklang mit der Konditionalisierungsregel revidiert, kein Dutch Book gegen ihn möglich ist. Mit anderen Worten, ein dynamisches Dutch Book gibt es genau dann nicht, wenn der Wetter seine Überzeugungen gemäß der einfachen Konditionalisierung revidiert. Diese Idee geht auf David Lewis Anfang der 70er Jahre zurück.[64]

Auch hier soll in einem kurzen (überspringbaren) Exkurs der relevante Teil des dynamischen Dutch Book-Theorems bewiesen werden. Das Theorem lautet:

> Wenn ein Akteur seine Überzeugungen (repräsentiert durch ein Wahrscheinlichkeitsmaß) nicht nach der einfachen Konditionalisierungsregel verändert, ist er einem dynamischen Dutch Book ausgesetzt.

Angenommen, der Wetter hat zu t ein altes Wahrscheinlichkeitsmaß P und macht zu t' die Erfahrung E aus einer Menge $\{E_1,...,E_n\}$, die eine Partition möglicher Erfahrungen darstellt, die alle eine positive Apriori-Wahrscheinlichkeit haben. Nach der Konditionalisierungsregel sollte er zu einem P' mit $P'(A) = P(A|E)$ für alle Propositionen A übergehen. Er verletzt die Regel, wenn es einen Fall gibt, wo $P'(A) \neq P(A|E)$. Betrachten wir den Fall, wo $P'(A) > P(A|E)$; der Teil des Beweises für $P'(A) < P(A|E)$ verläuft analog. Wenn also das neue Wahrscheinlichkeitsmaß unseres Wetters so aussieht, daß für ein A $P'(A) > P(A|E)$, so gibt es eine Zahl e, so daß $P'(A) - P(A|E) = e$. Nun wird der Akteur zu t zugleich folgende Wetten akzeptieren: (i) eine bedingte Wette derart, daß er im Fall von $A \wedge E$ dem Buchmacher $P(\neg A|E)$DM zahlt, im Fall von $\neg A \wedge E$ $P(A|E)$DM erhält, während im Fall von $\neg E$ keiner etwas zahlt oder gewinnt; (ii) eine Wette auf $\neg E$, so daß der Wetter im Falle von E $e \cdot P(\neg E)$DM gewinnt, im Falle von $\neg E$ hingegen $e \cdot P(E)$DM verliert.[65] Wenn nun zu t' $\neg E$ auftritt, so tritt Wette (i) nicht ein und der Akteur verliert nach Wette (ii) $e \cdot P(E)$DM. Wenn zu t' hingegen E eintritt, so wird der Akteur unter den vorausgesetzten Annahmen zu t' fol-

---

[64] Lewis' Beweis wird präsentiert in Teller (1976). Für eine detaillierte Darstellung des Arguments für verallgemeinerte Konditionalisierung cf. Skyrms (1987); diesem Aufsatz ist auch nachfolgender Beweis des Theorems für einfache Konditionalisierung entnommen. Cf. aber auch Earman (1992), der auf einige der Probleme mit Dutch-Book-Rechtfertigungen für Konditionalisierung hinweist.

[65] Zur Erinnerung: die Wahrscheinlichkeiten entsprechen den Wettquotienten, zu denen der Akteur bereit ist, auf die Wahrheit der jeweiligen Proposition zu wetten. Cf. zur Erläuterung die Diskussion des Dutch Book-Theorems im Kapitel 4.

gende Wette (iii) akzeptieren: wenn A eintritt, gewinnt der Akteur $P'(\neg A)$DM, wenn $\neg A$ der Fall sein wird, zahlt er $P'(A)$DM. Also erhält der Akteur im Falle $A \wedge E$ $P'(\neg A)$DM – $P(\neg A|E)$DM, welcher Betrag aber negativ ist, da $P'(\neg A) < P(\neg A|E)$ gdw. $P'(A) > P(A|E)$; und letzteres war vorausgesetzt. Wenn hingegen $\neg A \wedge E$ eintritt, so ergibt sich als „Gewinn" aus den Wetten (i) und (iii) $P(A|E)$DM – $P'(A)$DM, was unter der Voraussetzung sofort als negativ erkennbar ist. Ende des Exkurses.

Eine zweite Begründung der Konditionalisierung, die ich hier nicht ausführen will, findet sich ebenfalls in Teller (1976). Seine Grundidee ist, gewisse sehr schwach und plausibel aussehende Annahmen darüber zu machen, was während der doxastischen Änderung unverändert bleiben sollte, und dann zu beweisen, daß diese Annahmen effektiv so einschränkend sind, daß sie nur noch von der Konditionalisierungsregel erfüllt werden können.

Die dritte Rechtfertigungsstrategie geht schließlich von der Idee aus, daß die im Prozeß des Lernens aufgrund neuer Erfahrungen erfolgenden doxastischen Änderungen in gewisser Weise minimal ausfallen sollten; das ist eine Art Konservativitätsprinzip. Man sucht also nach einem neuen Wahrscheinlichkeitsmaß, das unter all den die gemachte Erfahrung inkorporierenden Wahrscheinlichkeitsmaßen dasjenige ist, welches den geringstmöglichen Abstand zum alten Maß aufweist. Die seltsame Rede vom Abstand zwischen zwei Wahrscheinlichkeitsmaßen wird dabei anhand von Ideen aus der Informationstheorie präzise erläutert und so ihres metaphorischen Charakters entkleidet; die daraus resultierende Änderungsregel heißt das Prinzip von der Minimierung der relativen Entropie. Man kann dann den absoluten und relativen Gehalt von Wahrscheinlichkeitsmaßen erfassen und zeigen, daß Konditionalisierung in der Tat eine minimale Änderung in diesem Sinne darstellt.[66]

Auch im Falle der dynamischen Glaubenstheorie gibt es weitere interessante und diskussionswürdige Rationalitätspostulate. Da gibt es z.B. das Reichenbach-Axiom. Es werde ein unfairer Würfel wiederholt geworfen. Das Axiom fordert nun, daß sich die subjektiven Wahrscheinlichkeiten den relativen Häufigkeiten angleichen sollten, die sich im Laufe des Würfelns herausbilden, und im Grenzwert mit ihnen zusammenfallen. Das Axiom ist ersichtlich ein Ausdruck von Lernfähigkeit. Ein noch einfacheres Axiom ist das Axiom der positiven Relevanz; es besagt, daß die Tatsache, daß ein Gegenstand eine gewisse Eigenschaft hat, es wahrscheinlicher macht, daß der

[66] Cf. hierzu Hunter (1991). Hunters Aufsatz enthält auch weitere Literaturverweise zu diesem Minimierungsprinzip.

nächste Gegenstand wieder diese Eigenschaft hat. Es läßt sich dann zeigen, daß das Reichenbach-Axiom aus diesem Axiom der positiven Relevanz und dem im Kapitel 4 erwähnten Symmetrie-Postulat herleitbar ist.[67] Möglicherweise das grundlegendste Prinzip der doxastischen Änderung ist van Fraassens Reflexionsprinzip.[68] Und so weiter; hier existiert ein insgesamt ungemein spannendes Theoriegut, welches auszubreiten viel Platz und auch nicht wenig Mathematik erforderte, so daß ich es bei den Verweisen belassen will.

In der jüngsten Zeit kommen einige der interessantesten Beiträge zur Theorie dynamischer Rationalität aus der KI-Forschung. Besonders zu erwähnen ist hier die Theorie Bayes'scher Netze, die klassische philosophische Konzeptionen mit neueren technischen Entwicklungen verbindet und die ich hier noch ein wenig ausführen will.[69] Genauer gesagt, habe ich bisher zwei Anknüpfungspunkte dafür ausgelegt. Im Kapitel 2 hatte ich Humes Vorstellung von einer Parallelführung kausalen und induktiven Räsonnierens dargelegt; die Bayes'schen Netze führen genau diese Vorstellung präzise aus. Und im Kapitel 4 hatte ich erwähnt, daß ein Wahrscheinlichkeitsmaß leicht gewissermaßen zu groß wird, um in einen Computer hineinzupassen; dieses Problem verbessert sich mit Hilfe der Bayes'schen Netze drastisch.

Um das näher zu erläutern, betrachten wir nochmals das Diagnosesystem für die ärztliche Praxis, dessen Design den KI-Forschern soviel Anstrengungen abforderte. Wir nehmen der Einfachheit halber an, daß alle relevante Variablen binärer Natur sind – ob also der Patient an einer Bronchitis leidet oder nicht, ob er einen Ohnmachtsanfall hatte oder nicht, ob er ein bestimmtes Medikament eingenommen hat oder nicht, usw. Wenn wir 100 derartige binäre Variablen haben – was nicht viele sind –, umfaßt der Möglichkeitsraum, wie wir gesehen hatten, 2^{100} mögliche Fälle. (Somit gibt es $2^{2^{100}}$ Propositionen über ihm!) Ein Wahrscheinlichkeitsmaß für ihn hatte also ohne einschränkende Bedingungen $2^{100} - 1$ Freiheitsgrade – was bedeutet, daß wir $2^{100} - 1$ Wahrscheinlichkeiten (innerhalb gewisser Grenzen) frei wählen können, und dann erst die letzte festliegt. Mit solchen Größenordnungen wird, wie gesagt, kein Computer je zurechtkommen.

Man könnte hoffen, die Freiheitsgrade durch das folgende Manöver zu reduzieren: Ein möglicher Fall besteht ja aus je einer Realisierung aller 100 Variablen, hat also

---

[67] Das alles läßt sich ausführlicher nachlesen in Stegmüller (1973).

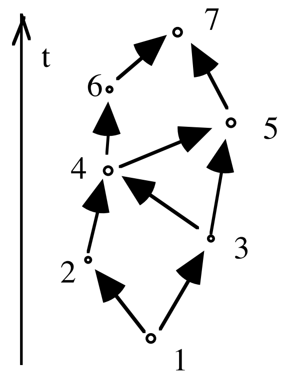
[68] Cf. van Fraassen (1984).

[69] Die Quelle schlechthin für diesen Ansatz ist Pearl (1988).

die Form $A_1 \wedge A_2 \wedge ... \wedge A_{100}$, worin A_k jeweils für eine Realisierung der Variablen k stehe. Dafür direkt die Wahrscheinlichkeit $P(A_1 \wedge A_2 \wedge ... \wedge A_{100})$ zu spezifizieren, ist ohnehin schwierig. Daher liegt es nahe, sie übers Multiplikationsprinzip (W9) in $P(A_1) \cdot P(A_2|A_1) \cdot \; ... \; \cdot P(A_{100}|A_1 \wedge A_2 \wedge ... \wedge A_{99})$ aufzulösen; dieses Produkt besteht aus lauter leichter einschätzbaren Faktoren. Reduziert sich damit die Anzahl der Freiheitsgrade? Leider nein. Mit $P(A_1)$ liegt bereits $P(\neg A_1)$ fest; aber dann ist sowohl $P(A_2|A_1)$ wie $P(A_2|\neg A_1)$ zu spezifizieren; und am Ende sind wieder 2^{99} Realisierungsweisen der ersten 99 Variablen als Bedingungen für den letzten Faktor des obigen Produkts zu betrachten. Das macht in summa wieder $2^{100} - 1$ Freiheitsgrade für die Festlegung eines Wahrscheinlichkeitsmaßes.

Trotzdem war die Überlegung hilfreich. Denn nun rückt die folgende Vereinfachung in den Blick: In aller Regel wird die Wahrscheinlichkeit von A_{100} ja nicht von allen anderen 99 Variablen abhängen! Wenn es aber nur vergleichsweise wenige Abhängigkeiten unter den einzelnen Variablen gibt, so reduziert sich die Zahl der zu berücksichtigenden Möglichkeiten doch ganz erheblich. Nehmen wir etwa an, A_{100} hänge nur von den vier Variablen 96 bis 99 ab; d.h. für die bedingte Wahrscheinlichkeit von A_{100} ist nur jeweils eine Realisierung dieser vier Variablen anzunehmen; sie bleibt dann immer dieselbe, wie immer sich die ersten 95 Variablen realisieren. Damit reduziert sich die Anzahl der Möglichkeiten für den letzten Faktor des obigen Produkts von 2^{99} auf $2^4 = 16$. Gilt das ebenso für alle anderen Variablen, so müssen wir für 100 Variablen jeweils 16 und damit insgesamt nur noch 1600 Freiheitsgrade berücksichtigen – eine enorme Reduktion, die sowohl für Computer wie auch für menschliche Gehirne handhabbar ist. Solche Abhängigkeitsannahmen bilden also hoch effektive Beschränkungen für mögliche Wahrscheinlichkeitsmaße.

Die Theorie Bayes'scher Netze liefert nun gerade ein Modell für die Analyse solcher Abhängigkeitsbeziehungen zwischen einzelnen Variablen. Am besten verdeutlicht man das Modell mit Hilfe einer Graphik:



Die mit Zahlen markierten Punkte stellen die Variablen dar. Wie der lange Pfeil links andeutet, sind sie in der Reihenfolge ihrer zeitlichen Realisierung angeordnet. Die kurzen Pfeile zwischen den Variablen drücken ihre Abhängigkeiten untereinander aus. Z.B. hängt die Variable 6 direkt nur von der Variablen 4 ab; von allen anderen früheren Variablen ist sie, gegeben eine Realisierung der Variablen 4, unabhängig. Das ganze Gebilde ist eine Repräsentation eines Bayes'schen Netzes. Aus der Graphik läßt sich nun direkt die Wahrscheinlichkeit von $A_1 \wedge A_2 \wedge A_3 \wedge A_4 \wedge A_5 \wedge A_6 \wedge A_7$ ablesen; sie ist:

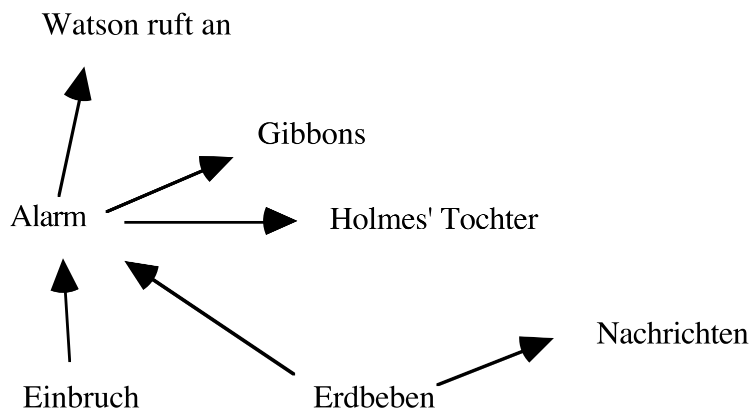
$$\begin{aligned} P(A_1 \wedge A_2 \wedge A_3 \wedge A_4 \wedge A_5 \wedge A_6 \wedge A_7) = \; & P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1) \cdot \\ & P(A_4|A_2 \wedge A_3) \cdot P(A_5|A_3 \wedge A_4) \cdot \\ & P(A_6|A_4) \cdot P(A_7|A_5 \wedge A_6). \end{aligned}$$

Hier vereinfacht sich etwa der dritte Faktor, weil aufgrund der durch die Graphik gegebenen Abhängigkeitsannahmen $P(A_3|A_1 \wedge A_2) = P(A_3|A_1 \wedge \neg A_2) = P(A_3|A_1)$ gilt. Diese probabilistischen Abhängigkeits- und Unabhängigkeitsbeziehungen, die hier nur graphisch dargestellt sind, werden nun in der Theorie Bayes'scher Netze in einer allgemeinen mathematischen Theorie erfaßt. Das ist schließlich vor allem deswegen bedeutsam, weil sich im Rahmen von statistischen Analysen der Kausalrelation die durch die Pfeile markierten probabilistischen Abhängigkeiten zwischen einzelnen Variablen als direkte kausale Abhängigkeiten verstehen lassen.[70]

Pearl (1988) gibt ein Beispiel für ein Bayes'sches Netz an, das direkt mit solchen Kausalbeziehungen operiert. Da ruft Watson Sherlock Holmes an und teilt ihm mit, in seinem Landhaus sei die Alarmanlage losgegangen. Holmes erschrickt und erwägt, hinauszufahren und nachzusehen, ob in seinem Landhaus eingebrochen worden ist.

[70] S. etwa Spohn (1994).

Andererseits weiß er, daß Watson zuweilen der Neigung nicht widerstehen kann, Holmes einen Streich zu spielen. Deshalb überlegt er als nächstes, seine Nachbarin auf dem Land, Mrs. Gibbons, anzurufen und zu fragen, ob sie auch den Alarm gehört habe. Da er annimmt, daß Mrs. Gibbons und Watson keinen Kontakt zueinander haben, sind ihre Auskünfte voneinander unabhängig. Bestätigt Mrs. Gibbons den Alarm, hat Holmes bessere Gründe, nach dem Rechten zu sehen. Das zeigt im Rahmen des Modells Bayes'scher Netze den Wert unabhängiger Zeugen. Sollte natürlich Watson Mrs. Gibbons in seinen Scherz einbezogen haben, so gäbe die Aussage von Mrs. Gibbons dem guten Sherlock Holmes über Watsons Anruf hinaus keinen weiteren Aufschluß über den Alarm. Aber Pearls Geschichte geht noch weiter. Mrs. Gibbons pflegt dem Alkohol über Gebühr zuzusprechen und hört und sieht in dem dann resultierenden Zustand kaum mehr etwas, weshalb ihren Aussagen wenig wert sind. Aber Holmes' Tochter sollte gleich ins Landhaus kommen, weswegen er nach den Gesprächen mit Watson und Mrs. Gibbons besser noch ihren Anruf abwartet. Außerdem denkt Holmes, daß der Alarm statt durch einen Einbruch durch ein schwaches Erdbeben ausgelöst worden sein könnte. Weil in England auch schwache Erdbeben sehr selten sind, würde das Auftreten eines solchen sicher im Rundfunk gemeldet; also schaltet Holmes BBC ein. Nachrichtensendungen sind im allgemeinen zuverlässige Indikatoren für Ereignisse, und daher wird Holmes' Wahrscheinlichkeit dafür, daß die Alarmanlage im Landhaus anging, steigen, wenn er in der BBC erfährt, daß just in dieser Gegend die Erde gebebt hat; das Beben wäre eine gemeinsame Ursache für die Nachricht wie für den Alarm. Das Bayes'sche Netz dieser Geschichte sieht also so aus:



Ein entscheidender Vorteil der Bayes'schen Netze besteht darin, daß man sofort sieht, wie sich die Information über Kausalzusammenhänge lokal im Netz fortpflanzt.

Das ermöglicht eine differenziertere Darstellung des Einflusses neuer Information, als es die verschiedenen Formen von Konditionalisierung erlaubten, die zunächst nur global sagten, wie Informationen Überzeugungssysteme verändern. Damit wird schließlich auch meine Behauptung verständlich, daß die Bayes'schen Netze gerade Humes noch ganz informelles Bild vom kausalen Schließen präzisieren.

Über die Dynamik probabilistisch repräsentierter Überzeugungen existieren also schon lange Vorstellungen (die freilich erst jüngst sorgfältig diskutiert werden). Die Dynamik qualitativ repräsentierter Überzeugungen ist hingegen erst in den letzten 15 Jahren überhaupt theoretisch erfaßt worden. Darunter ragt der Ansatz von Peter Gärdenfors hervor, den ich als nächsten darstellen will.[71]

Gärdenfors akzeptiert die im vorigen Kapitel erläuterten Gesetze (G1) - (G4) qualitativen Glaubens; er behandelt mithin Überzeugungen als propositionale Einstellungen im strikten Sinne und stellt so doxastische Zustände in idealisierter Weise als Mengen von Überzeugungen dar, die (G1) - (G4) genügen, d.h. konsistente und unter logischer Folgerung abgeschlossene Mengen von Propositionen sind.[72] Ein rationaler Denker kann also drei verschiedene doxastische Haltungen zu einer Proposition einnehmen: er kann sie akzeptieren, er kann sie zurückweisen oder er kann sich neutral dazu verhalten. Wenn er *A* akzeptiert, sagen wir, daß *A* Element seiner Überzeugungsmenge ist, wenn er *A* zurückweist, dann ist non-*A* Element seiner Überzeugungsmenge; ansonsten sind weder *A* noch non-*A* darin enthalten.

So weit die Rekapitulation der statischen Gegebenheiten. Gärdenfors' Untersuchungsgegenstand sind nun die Gesetze der Dynamik oder Änderung solcher Überzeugungsmengen. Grundsätzlich unterscheidet er zwischen drei Typen von Veränderung. Da gibt es den Fall der *Erweiterung* einer Überzeugungsmenge um eine Proposition *A*: das bedeutet, daß das Subjekt von der Neutralität zur Akzeptierung (oder Zurückweisung) von *A* übergeht. Der zweite Fall ist der der *Revision*: einer Überzeugungsmenge: hier ändert das Subjekt seine Einstellung zu mindestens einer Proposition *A* von der Zurückweisung zur Akzeptierung (oder umgekehrt). Schließlich haben wir noch den Fall der *Kontraktion*: ein Subjekt hat bislang eine Proposition *A* akzep-

---

[71] Für das folgende ist nachdrücklich auf Gärdenfors (1988) zu verweisen, wo eine sehr klare Darstellung der Theorie und ihrer Anwendungen zu finden ist. Zudem setzt Gärdenfors seinen Ansatz auch in Beziehung zu verschiedenen anderen, insbesondere probabilistischen Modellen rationaler Überzeugungsänderung.

[72] Das heißt, wenn K eine solche Überzeugungsmenge ist, gilt: (1) für alle A nicht zugleich $\neg A \in K$ und $A \in K$; (2) für alle B, die aus A logisch folgen, gilt mit $A \in K$ auch $B \in K$.

tiert (oder zurückgewiesen) und verändert seine Einstellung zu *A* zur Neutralität. Für alle diese Typen von Änderung gilt es nun Gesetze anzugeben.

Die Erweiterung einer Überzeugungsmenge ist der einfachste Fall. Sie tritt ein, wenn man etwas Neues erfährt, was mit der bisherigen Überzeugungsmenge verträglich ist. Ich sehe vor meinem Bürofenster einen Turmfalken vorbeifliegen; darauf füge ich meiner Überzeugungsmenge hinzu, daß es rund um die Uni Bielefeld Turmfalken gibt. Gärdenfors gibt, in Form von Axiomen, Bedingungen an, denen der Prozeß der Erweiterung einer Überzeugungsmenge seiner Ansicht nach genügen muß. Dabei stehe *K* im folgenden für die Überzeugungsmenge vor der Änderung. Die Erweiterung von *K* um *A* sei mit K_A^+ bezeichnet; wobei vorausgesetzt sei, daß non-*A* nicht in *K* enthalten ist. + ist hier als eine Funktion aufzufassen, die Paare aus Überzeugungsmengen und Propositionen in Überzeugungsmengen abbildet.[73] Gärdenfors' Axiome für Erweiterung sehen nun so aus:

(K^+1) K_A^+ ist für alle *K* und *A* eine Überzeugungsmenge, d.h. konsistent und deduktiv abgeschlossen,
(K^+2) $A \in K_A^+$,
(K^+3) $K \subseteq K_A^+$,
(K^+4) Wenn $A \in K$, dann $K_A^+ = K$,
(K^+5) Wenn $K \subseteq H$, dann $K_A^+ \subseteq H_A^+$,
(K^+6) Für alle Überzeugungsmengen *K* und Sätze *A* gilt: K_A^+ ist die kleinste Überzeugungsmenge, die (K^+1)–(K^+5) erfüllt.

Ein paar Erläuterungen dazu: die beiden ersten Axiome verstehen sich eigentlich von selbst; das dritte Axiom zeigt an, warum der Prozeß eine Erweiterung einer vorgegebenen Überzeugungsmenge ist. Das vierte Axiom ist klar: wenn eine Proposition *A* schon in *K* ist, dann passiert bei der Erweiterung um *A* nichts. Kritischer ist (K^+5): es besagt, daß die Überzeugungsmenge *H*, sofern sie vor der Änderung reichhaltiger ist als *K*, auch nach der Hinzufügung von *A* zu beiden Mengen reichhaltiger bleibt; es heißt daher ein Monotonie-Axiom. Das sechste Axiom ist Ausdruck der Forderung, daß K_A^+ neben den von *A* erzwungenen keine zufälligen Anreicherungen enthalten sollte; es ist also wieder eine Art Minimalitätsprinzip für Änderungen. Auf der Basis dieser Axiome kann man nun beweisen, daß die Funktion '+' die Axiome (K^+1) bis

[73] Gärdenfors redet durchweg von Sätzen statt von Propositionen; da er aber (G4) akzeptiert, ist meine Darstellung mit der seinen äquivalent.

(K^+6) genau dann erfüllt, wenn K_A^+ gleich der Menge aller Propositionen ist, die aus K und A logisch folgen (Beweis bei Gärdenfors 1988, Appendix A).

Wesentlich komplizierter, aber auch interessanter wird das Bild, wenn man Revisionen untersucht, die wir ja ebenfalls laufend vollziehen. Bisher habe ich geglaubt, daß das Bielefelder Unigebäude keine Brutplätze für Vögel bietet, so unwirtlich glatt, wie der Beton ist. Aber nun sehe ich, wie die Turmfalken erbeutete Mäuse zu den Aufzugsbauten oben auf dem Dach der Uni bringen, wo die Jungen auf Nahrung warten. Also werde ich meine erstgenannte Überzeugung revidieren. Der Fall ist diffiziler, weil sich die Revision nicht auf diese eine Überzeugung beschränken kann. Und weil der Prozeß nicht eindeutig ist: vielleicht können die Jungen entgegen anderen Überzeugungen, die ich bisher hatte, schon fliegen und sind da oben hinaufgeflogen. All dies hat unter anderem zur Konsequenz, daß die Monotonie, die für Erweiterungen zutreffend war, im Fall von Revisionen nicht gewahrt bleiben muß. Formal gesehen, betrachtet Gärdenfors auch die Revision als eine Funktion $*$, die Paare aus Überzeugungsmengen und Propositionen in Überzeugungsmengen abbildet; K_A^* stehe also für die Revision der Überzeugungsmenge K im Hinblick auf die neue Proposition A. Die beiden ersten von Gärdenfors' Axiomen für die Revision sind wiederum klar:

(K*1) K_A^* ist für alle K und A eine Überzeugungsmenge,

(K*2) $A \in K_A^*$

.

Ferner ist es plausibel anzunehmen, daß sich die Revision bzgl. A in dem Fall, in dem K mit A verträglich ist, einfach auf die Erweiterung um A reduziert:

(K*3) Wenn $\neg A \notin K$, dann $K_A^* = K_A^+$.

Damit erweist sich die Erweiterung formal als ein Spezialfall der Revision.

(K*1) - (K*3) bilden den minimalen Kern der Revisionsaxiome. Die Diskussion um Verstärkungen ist mittlerweile zu umfangreich und detailliert geworden, um sich hier darstellen zu lassen. Erwähnt sei nur noch eines der stärksten Axiome, welches iterierte Revisionen betrifft und insofern reich an Konsequenzen ist. Es lautet:

(K*4) wenn $\neg B \notin K_A^*$, dann $K_{A \cap B}^* = (K_A^*)_B^+$.

Umgangssprachlich besagt dies, daß es auf dasselbe hinausläuft, ob man erst bzgl. A revidiert und dann um das mit dieser Revision verträgliche B oder ob man gleich bzgl. der gesamten Information $A \wedge B$ revidiert. Dies bedeutet eine erhebliche Einschränkung für die Hintereinanderschaltung mehrerer Überzeugungsänderungen.

Es ist aber insgesamt festzuhalten, daß diese Axiome für sich genommen nicht dazu ausreichen, die genaue Art der Revision festzulegen, die die Hinzufügung einer Proposition A zu eine Menge K erzwingt. Im allgemeinen gibt es viele verschiedene Möglichkeiten, die Konsistenz der Überzeugungsmenge K_A^* herzustellen; die Axiome legen dafür nur einen Rahmen fest. D.h. es gibt im allgemeinen viele Funktionen $*$, die diesen Axiomen genügen. Diese sind einer systematischen Beschreibung zugeführt worden, wozu ich nach der Darstellung von Kontraktionen noch kurz einige Worte verlieren will.

Kontraktionen stellen ebenfalls eine wichtige Form von Überzeugungsänderung dar. Wenn ich überlege, ob im Bereich der Uni Bielefeld anstelle der Turmfalken auch seltenere Falkenarten leben könnten, dann klammere ich gleichsam einige meiner Überzeugungen, die tatsächlich meiner Überzeugungsmenge angehören, ein und überlege, welche Änderungen dann in meiner Überzeugungsmenge auftreten können. Wichtig ist in diesem Zusammenhang ein Prinzip der Wahrung von Information: wenn A aus K entfernt wird, muß zwar so viel mit entfernt werden, daß aus dem verbleibenden Rest A nicht mehr folgt; es darf aber nicht willkürlich mehr hinausgeworfen werden. Wie im Falle der Revision nimmt Gärdenfors an, daß sich Kontraktionen durch eine Funktion, mitgeteilt als ‚–', darstellen lassen, deren Argumente Überzeugungsmengen K und Propositionen A sind und deren Werte Überzeugungsmengen K_A^- sind. Die ersten der Axiome sind wieder unproblematisch:

(K^-1) K_A^- ist für alle K und A eine Überzeugungsmenge,
(K^-2) $K_A^- \subseteq K$,
(K^-3) wenn $A \notin K$, dann $K_A^- = K$.

Denn es wird bei einer Kontraktion ja mindestens A aus K entfernt und wenn A gar nicht in K ist, so soll sich nichts ändern. Klar ist auch, daß eine Kontraktion erfolgreich sein soll, daß also A nicht eine logische Folgerung aus K_A^- ist, weshalb gefordert wird:

(K^-4) Wenn A nicht die tautologische Proposition ist, so $A \notin K_A^-$.

Kniffliger ist das Prinzip, daß eine Kontraktion um A und eine anschließende Erweiterung um A an K nichts ändern soll. Gärdenfors fordert nur die eine Richtung in Gestalt von

$$(K^-5) \qquad \text{wenn } A \in K, \text{ so } K \subseteq (K_A^-)_A^+$$

als Axiom; die Umkehrung „wenn $A \in K$, so $(K_A^-)_A^+ \subseteq K$“ ergibt sich als logische Folgerung der ersten vier Axiome für Kontraktionen (Beweis bei Gärdenfors 1988, Appendix A). Jeder teste für sich selbst, ob eine Entfernung einer Überzeugung aus seiner Glaubensmenge und ihre neuerliche Hinzufügung den vorangegangenen Zustand wieder herstellt oder nicht.[74] Wie sich allerdings später zeigt, ist dieses Axiom für die fundamentalen Charakterisierungstheoreme nicht erforderlich.

Wie für Revisionen gilt auch hier, daß die vorgestellten Axiome allein noch nicht ausreichen, um Kontraktionen eindeutig zu charakterisieren. Und wiederum haben Gärdenfors und viele mit ihm sinnvolle Verstärkungen dieser Axiome diskutiert. Man wird schon jetzt den Verdacht hegen, daß Revisionen und Kontraktionen eng miteinander zusammenhängen; tatsächlich lassen sich die Verfahren wechselseitig durcheinander definieren. Nach den Philosophen, die erstmals darauf hingewiesen haben, nennt man die Gesetze

(Def*) $K_A^* = (K_{\neg A}^-)_A^+$ die Levi-Identität, und
(Def–) $K_A^- = K \cap K_{\neg A}^*$ die Harper-Identität.

Mit der Hilfe dieser beiden Identitäten erweist sich, daß die axiomatischen Charakterisierungen von Revisionen und Kontraktionen einander entsprechen. Das heißt, es gilt, wie Gärdenfors beweist, unter anderem folgendes: Sei die Revisionsfunktion $*$ mittels der Levi-Identität aus der Kontraktionsfunktion – definiert. Wenn nun die Kontraktionsfunktion (K^-1) - (K^-4) und die Erweiterungsfunktion + (K^+1)-(K^+3) erfüllt, dann erfüllt die Revisionsfunktion (K^*1)-(K^*3). Und umgekehrt: wenn die Revisionsfunktion $*$ (K^*1)-(K^*3) erfüllt, dann erfüllt die über die Harper-Identität gewonnene Kontraktionsfunktion – (K^-1)-(K^-5).[75]

---

[74] Es ist übrigens im allgemeinen nicht der Fall, daß $(K_A^+)_A^- = K$, wie Gärdenfors (1988) ausführt.

[75] Cf. Gärdenfors (1988) für die Beweise. Man beachte allerdings, daß bei der Charakterisierung der über (Def*) erhaltenen Repräsentation das Axiom (K^-5) nicht verwendet wird.

Neben dieser axiomatischen Charakterisierung von Überzeugungsänderungen gibt Gärdenfors auch andere Charakterisierungen an. Wie die zuletzt angestellten Überlegungen zeigen, lassen sich alle Änderungen auf einen Typ zurückführen, von dem aus die anderen dann definiert werden können. Fassen wir etwa Kontraktionen als grundlegend auf. Gärdenfors (1988), Kap. 4, gibt dann zwei im wesentlichen äquivalente Konstruktionsverfahren für Kontraktionen an. Das erste nimmt seinen Ausgang von sogenannten minimalen Kontraktionen: minimal ist eine Kontraktion um A genau dann, wenn die daraus resultierende Menge K_A^- in dem Sinne maximal ist, daß zu K_A^- nichts hinzugefügt werden kann, ohne daß A wieder daraus folgte. Solche minimalen Kontraktionen haben Eigenschaften, die intuitiv nicht erwünscht sind. Doch lassen sich alle interessanten Kontraktionen, die den obigen Axiomen genügen, als Durchschnitte solcher minimalen Kontraktionen konstruieren.

Das zweite Konstruktionsverfahren basiert darauf, daß sich die Überzeugungen in einer Überzeugungsmenge nach dem Grad ihrer doxastischen Verankerung ordnen lassen. Die doxastische Verankerung einer Überzeugung besteht nicht unbedingt in hoher Wahrscheinlichkeit, sondern in der Wichtigkeit dieser Überzeugung innerhalb eines vorgegebenen Korpus von Überzeugungen. Als Illustration diene etwa die eines Physikers, in der gesetzesartige Aussagen im allgemeinen einen höheren Grad an doxastischer Verankerung besitzen als kontingente Einzelaussagen – was sich gerade darin äußert, daß diese gesetzesartigen Aussagen weniger schnell einer doxastischen Änderung zum Opfer fallen als schlechter verankerte. Gärdenfors fordert nun von einer Relation der komparativen doxastischen Verankerung bestimmte plausible Ordnungseigenschaften und konstruiert dann aus einer solchen Ordnungsrelationen die entsprechenden Kontraktionsfunktion – wonach, grob gesagt, bei einer Kontraktion um A zunächst die am schlechtesten verankerten Propositionen aufgegeben werden, dann die am zweitschlechtesten verankerten, und so weiter, bis A aus dem verbleibenden Rest nicht mehr folgt. Für eine genaue Darstellung all dieser Zusammenhänge kann ich abr nur auf Gärdenfors' Buch verweisen.

Wie man an der Idee einer Ordnung doxastischer Verankerung schon sieht, spielen Grade der Überzeugtheit in der Dynamik von Glauben auch im qualitativen Rahmen eine Rolle – auch wenn sie sich in Gärdenfors' Rahmen nicht mit Wahrscheinlichkeiten in Entsprechung bringen lassen.[76] Der bisher betrachtete qualitative Ansatz hat freilich verschiedene Nachteile, die sich im Kern darauf zurückführen lassen, daß

[76] Gärdenfors (1988) diskutiert allerdings auch probabilistische Modelle und führt in diesem Rahmen für Wahrscheinlichkeitsmaße analoge Prozesse der Kontraktion und Revision ein.

iterierte Überzeugungsänderungen auch mit dem Axiom (K*4) noch sehr unzulänglich erfaßt sind.

Es gibt aber noch einen weiteren Ansatz zur Modellierung von Überzeugungsänderungen, der diese und die daraus resultierenden Lücken zu schließen imstande ist. Bei diesem Ansatz handelt es sich um die Theorie der sogenannten *Rangfunktionen*, die ich selbst entwickelt habe und die insbesondere von dem schon erwähnten Judea Pearl und seinen Mitarbeitern aufgegriffen wurde - wohl auch deswegen, weil diese Rangfunktionen sich in vielem ganz ähnlich wie Wahrscheinlichkeitsmaße verhalten und insbesondere eine strikte Analogie zu den Bayes'schen Netzen zu konstruieren gestatten.[77]

In dieser Theorie werden doxastische Zustände durch Rangfunktionen repräsentiert. Eine Rangfunktion weist dabei jeder Proposition einen Rang, d.h. eine natürliche Zahl (bzw. in einer allgemeineren Version eine Ordinalzahl) zu, die ihren Grad des Für-falsch-gehalten-Werdens zum Ausdruck bringt. Der Rang 0 besagt, daß die Proposition gar nicht für falsch gehalten wird (was nicht unbedingt bedeutet, daß sie für wahr gehalten wird); und die Ränge 1,2,... stehen für immer stärkere Grade des Für-falsch-Haltens. Für eine formale Definition gehen wir wieder von einem Möglichkeitsraum Ω aus, der aus einzelnen möglichen Fällen oder Möglichkeiten w, w', etc. besteht; und Propositionen sind wiederum als Teilmengen von Ω dargestellt. Damit sei definiert:

> κ ist eine Rangfunktion gdw. κ eine Funktion von Ω in die Menge der natürlichen Zahlen ist derart, daß (a) $\kappa^{-1}(0) \neq \emptyset$, d.h. daß $\kappa(w) = 0$ für wenigstens ein $w \in \Omega$. Auf Propositionen sei diese Funktion durch die Festlegung (b) $\kappa(A) = \min\{\kappa(w) | \kappa \in A\}$ für alle $A \neq \emptyset$ ausgedehnt; und schließlich sei (c) $\kappa(\emptyset) = \infty$.

Der Grad, in dem eine Proposition für falsch gehalten wird, kann natürlich nicht größer sein als der Grad, in dem die in ihr enthaltenen Möglichkeiten für falsch gehalten werden. Daß er aber auch nicht kleiner ist, ist Inhalt der Festlegung (b). Die Bedingung (a) stellt sicher, daß es Möglichkeiten gibt, die gar nicht für falsch gehalten

[77] Der zentrale Aufsatz für diese Konzeption ist Spohn (1988). Diese Arbeit ist aber sehr technisch; eine zugänglichere als Präsentation findet sich in Spohn (1991), dem das folgende weitgehend entnommen ist. „Rangfunktion", englisch „ranking function" ist ein Terminus, den Judea Pearl eingeführt hat und den ich mittlerweile bevorzuge; in meinen Originalschriften sprach ich noch von ordinalen und naürlichen Konditionalfunktionen.

werden; denn würden alle Möglichkeiten für unzutreffend erachtet, so nach (b) auch die tautologische Proposition Ω, was absurd wäre.[78] Die Bedingung (c) besagt zuletzt, daß die kontradiktorische Proposition in stärkerem Maße für falsch gehalten wird als alle anderen Propositionen; der Wert ∞ darf keiner anderen Proposition zugewiesen werden. Das entspricht dem im letzten Kapitel erwähnten Regularitätsaxiom in der Wahrscheinlichkeitstheorie.

Die Rede vom Für-falsch-Halten ist etwas umständlich, gibt aber genau die intendierte Interpretation. Man beachte also, daß $\kappa(A) = 0$ nicht heißt, daß A im doxastischen Zustand κ geglaubt, akzeptiert oder für wahr gehalten wird; es heißt nur, daß A nicht für falsch gehalten wird. Daß A für wahr gehalten wird, heißt vielmehr, daß non-A für falsch gehalten wird, was durch $\kappa(\text{non-}A) > 0$ auszudrücken ist. In der Tat folgt, wie man sich leicht klar macht, aus der Definition für jede Proposition A, daß $\kappa(A) = 0$ oder $\kappa(\text{non-}A) = 0$; das heißt, man kann nicht sowohl A wie non-A für falsch halten. Aber es kann auch $\kappa(A) = 0$ und $\kappa(\text{non-}A) = 0$ gelten; das heißt, man kann sowohl A wie auch non-A nicht für wahr halten, d.h. sich gegenüber A neutral verhalten.

Um das etwas transparenter zu machen, empfiehlt es sich, für eine Rangfunktion κ die zugehörige Glaubensfunktion β zu definieren: β ist eine Funktion von der Menge der Propositionen in die um +∞ und –∞ erweiterte Menge der natürlichen Zahlen, für die gilt: $\beta(A) = \kappa(\text{non-}A) - \kappa(A)$. Dann kann man sagen: A wird genau dann geglaubt, wenn $\beta(A) > 0$, für falsch gehalten, wenn $\beta(A) < 0$, und neutral betrachtet, wenn $\beta(A) = 0$. Der β-Wert einer Proposition ergibt sich aber nicht so einfach aus den β-Werten der in ihr enthaltenen Möglichkeiten, wie das definitionsgemäß bei den κ-Werten der Fall war. Von daher empfiehlt es sich, mit den Rangfunktionen als dem grundlegenderen Begriff zu operieren.

So weit spielte die Möglichkeit, die Ränge, d.h. die Werte von κ bzw. β als *Grade* des Für-falsch- bzw. -wahr-Haltens zu interpretieren, noch keine Rolle. Ihre theoretische Relevanz erhalten sie erst, wenn wir die Dynamik der durch Rangfunktionen repräsentierten doxastischen Zustände betrachten. Dazu definieren wir zunächst für eine nicht-leere Proposition A den *A-Teil von κ*: das ist die auf A definierte Funktion $\kappa(\text{-}|A)$, für die $\kappa(w|A) = \kappa(w) - \kappa(A)$ für alle $w \in A$. Auf Propositionen erweitert wird

[78] κ^{-1} bezeichnet die Umkehrung der Funktion κ. Das muß nicht unbedingt wieder eine Funktion sein (und wird es im vorliegenden Fall auch meist nicht sein). Wenn die Funktion κ irgendwelchen Objekten aus ihrem Definitionsbereich den Wert 0 zuordnet, dann ordnet κ^{-1} dem Argument 0 als Wert alle die Objekte zu, denen κ den Wert 0 zugeordnet hatte.

der A-Teil von κ durch die Festlegung $\kappa(B|A) = \min\{\kappa(w|A) \mid w \in A \wedge B\} = \kappa(A \wedge B) - \kappa(A)$, für alle $B \subseteq \Omega$. $\kappa(B|A)$ drückt also aus, in welchem Grade B unter der Bedingung A für falsch gehalten wird; das ist genau die rangfunktionale Entsprechung der Definition bedingter Wahrscheinlichkeiten.

Mit diesem Hilfsbegriff läßt sich nun die Dynamik von Rangfunktionen beschreiben. Das Bild ist dies: der doxastische Ausgangszustand eines Subjekts wird zu einem bestimmten Zeitpunkt durch κ repräsentiert. Nun erhält das Subjekt die Information A, und zwar mit einem gewissen Stärke- oder Verläßlichkeitsgrad n. Das heißt, daß der neue Rang für non-A, d.h. der Grad, in dem non-A für falsch gehalten wird, gleich n ist. Das führt zu einem neuen doxastischen Zustand $\kappa_{A,\mathrm{n}}$ nach der Information, der wie folgt definiert ist: $\kappa_{A,n}(w) = \kappa(w|A)$ für den Fall, daß $w \in A$, und $\kappa_{A,n}(w) = n + \kappa(w|\text{non-}A)$ für den Fall, daß $w \in \text{non-}A$. Intuitiv hat das den folgenden plausiblen Inhalt: die Information, daß A der Fall ist, ändert am A-Teil wie auch am non-A-Teil der Rangfunktion κ nichts; neu ist nur der veränderte Rang von A. Und nur insoweit ändern sich die Ränge einzelner Möglichkeiten: die mit A verträglichen Möglichkeiten erhalten den Rang, den sie im alten A-Teil von κ hatten; und der Rang der mit A nicht verträglichen Möglichkeiten verschiebt sich gemäß dem Grad, in dem A nun geglaubt wird, ist also gleich diesem Grad plus dem Rang im alten non-A-Teil von κ. Diesen Vorgang bezeichne ich als A,n-Konditionalisierung einer Rangfunktion. Er beruht ersichtlich auf denselben Änderungsprinzipien wie die Jeffrey-Konditionalisierung und stellt das zentrale dynamische Gesetz der Theorie der Rangfunktionen dar.

Damit lassen sich, im Gegensatz zu Gärdenfors' Theorie, problemlos und uneingeschränkt iterierte Überzeugungsänderungen behandeln – was noch andere Vorteile nach sich zieht. Insgesamt besteht eine starke Analogie zwischen der Theorie der Rangfunktionen und der Wahrscheinlichkeitstheorie, die sich systematisch nachweisen ließe; insbesondere erstreckt sich diese Analogie auch auf die oben angerissene Theorie der Bayes'schen Netze. Doch muß ich diesbezüglich auf die angegebene Literatur verweisen.[79] Generell läßt sich sagen, daß die Untersuchung der Zusammenhänge zwischen all diesen erst in den letzten Jahren aufgekommenen Repräsentationen doxastischer Zustände und ihrer rationalen Änderungen einen der derzeit aktivsten und spannendsten Bereiche der Erkenntnistheorie darstellt, der sicherlich noch viele interessante Ergebnisse birgt.

---

[79] Zu Zusammenhängen zwischen den Theorien von Spohn und von Gärdenfors vgl. auch Gärdenfors (1988).

Nach diesen ausgiebigen Diskussionen rationaler Prinzipien des Glaubens und seiner Dynamik erlaube ich mir noch eine kleine kritische Randbemerkung mit Blick auf die philosophische Szene in Deutschland: In vielen Diskussionen, die hierzulande laufen, spielt der Begriff der Rationalität eine zentrale Rolle. Man liest so allerhand über eine Krise der Rationalität, über unvollständige Konzeptionen von Rationalität, die es zu erweitern oder zu ändern gelte – und bei alledem ist leider festzustellen, daß der angesprochene Begriff der Rationalität zumeist höchst unklar und schwammig ist. Nicht der geringste Zweck des bis hierhin Ausgeführten besteht darin zu zeigen, daß es zumindest im Bereich der Erkenntnistheorie eine substantielle und reichhaltige Konzeption epistemischer Rationalität gibt, deren Kenntnisnahme für so manche philosophische Debatte kein Nachteil wäre.

6. Gründe und Begründungen

Nach dieser ausgiebigen Diskussion der Frage, wie man Überzeugungen und ihre rationale Veränderung angemessen darstellt, wird sich vielleicht manch einer schon ungeduldig die Frage gestellt haben, ob es nicht eine viel einfachere Antwort auf die Frage, was die Rationalität einer Meinung oder ihrer Änderung ausmacht, gegeben hätte: Eine Meinung oder ihre Änderung ist genau dann rational, wenn sie gut begründet ist. Das scheint viel weniger aufwendig als all die Betrachtungen über Dutch Books, Konditionalisierungsregeln etc. Freilich ist nicht von vornherein klar, daß diese Antwort eine Alternative zu der vorangegangenen Diskussion in sich birgt. Zwar haben wir, wenn man so will, bisher nur formale Kriterien doxastischer Rationalität studiert, während der Verweis auf gute Gründe eine substantiellere Rationalitätstheorie zu versprechen scheint. Doch gilt es dazu erst einmal, den Begriff des Grundes oder der Begründung genauer zu verstehen. Dies will ich dadurch leisten, daß ich zunächst fünf Bedingungen erläutere, denen der Begründungsbegriff plausiblerweise gerecht werden sollte, und dann vier verschiedene Begründungsbegriffe vorstelle, die in der Literatur prominent sind, und überprüfe, wie sie sich zu den fünf Bedingungen verhalten. Die Schlußfolgerung wird sein, daß der relativ brauchbarste Begriff des Grundes gerade auf die im letzten Kapitel geführte Diskussion Bezug nimmt – so daß sich die von mir gewählte Reihenfolge der Themen nicht als verkehrt, sondern als konsequent erweist.

In seinem jüngsten Buch (1993) hat Robert Nozick auf zwei Aspekte hingewiesen, die jede Theorie der Rationalität berücksichtigen muß. Der eine Aspekt ist, daß die Rationalität einer Überzeugung von den Gründen für sie abhängt; was das bedeutet, wollen wir in diesem Kapitel herausfinden. Der andere Aspekt ist, daß Rationalität etwas mit der Zuverlässigkeit („reliability“) der Prozesse zu tun, die Überzeugungen hervorrufen; rationale Überzeugungsbildung ist nach Nozick auch dadurch gekennzeichnet, daß sie Überzeugungen zeitigt, die vielleicht nicht mit absoluter Sicherheit, aber doch mit beträchtlicher Zuverlässigkeit, oder mit größerer Zuverlässigkeit als weniger rationale Methoden wahr oder zutreffend sind. Zusammengenommen bedeutet dies also, daß Gründe oder Begründungen zuverlässig sein müssen, d.h. mit relativer Zuverlässigkeit zu wahren Überzeugungen führen. Dies ist unsere erste Bedingung.

Die zweite Bedingung liegt in der Rede von guten Gründen, die ein evaluatives oder normatives Moment im Begründungsgeschäft andeutet. Wenn wir etwas als einen Grund akzeptieren, so bewerten wir ihn ipso facto als guten Grund; die Einsicht, ein gegebener Grund sei ein schlechter Grund, führt automatisch zu seiner Aufgabe. Natürlich kann man sich über die Bewertung streiten; und es kann einer die Gründe, die ein anderer für eine Meinung hat, für schlecht halten und dann mit geeigneten Argumenten des anderen Meinung zu unterminieren trachten. Aber in Bezug auf einen selbst geht das nicht. Dies ist also die zweite Bedingung: Gründe für Überzeugungen sind bewertbar und vom Subjekt selbst stets als gut bewertet.

Ferner ist festzustellen, daß es uns ganz natürlich ist, daß Gründe mehr oder weniger stark oder schwach sein können. Auch gibt es Gegengründe; ein Grund kann für oder wider eine bestimmte Meinung sprechen. Eine Begründungstheorie muß beidem Rechnung tragen – das ist die dritte Bedingung. Man darf hierbei die normative Unterscheidung zwischen guten und schlechten Gründen nicht mit der Unterscheidung zwischen starken und schwachen Gründen verwechseln; ein schwacher Grund kann sehr wohl als schwacher akzeptiert und insoweit als gut befunden sein.

Eine weitere wichtige Unterscheidung gilt es zu beachten: es ist eine Sache, zu sagen, daß *A* ein Grund für *B ist*, und eine andere, daß *S A* als Grund für *B hat*. Damit eine Meinung, die eine Person hat, begründet ist, muß es nicht bloß Gründe für sie geben, sondern sie muß diese Gründe auch tatsächlich haben, d.h. von ihrem Vorliegen überzeugt sein. Zum Beispiel gibt es viele Sachverhalte, die dafür sprechen oder Gründe dafür sind, daß ab morgen der universale Frieden auf der Welt ausbricht: etwa der Sachverhalt, daß morgen alle zu der wesentlich friedlichen Religion des Taoismus bekehrt werden; oder der, daß ab morgen alle Regierungen dieser Welt uneingeschränkt die Menschenrechte achten. Das sind ohne Zweifel Gründe für den morgigen Weltfrieden, die wir sogar als gute Gründe bewerten würden. Aber kein Mensch hat diese Gründe, kein Mensch glaubt an das Zutreffen dieser Gründe, und deswegen glaubt auch keiner an den morgigen Weltfrieden. Die vierte Bedingung ist also, daß eine Begründungstheorie beides leisten muß: eine Begründungsrelation zwischen möglichen Propositionen spezifizieren, unabhängig davon, ob man sie für wahr hält oder nicht, und sagen, wann man in einem doxastischen Zustand einen Grund für eine Überzeugung hat.

Schließlich – es klang im letzten Abschnitt schon an – sind Gründe selbst wieder Gegenstand rationaler Begründung. Damit droht jedoch das schon im Kapitel 3 eingeführte Begründungstrilemma: entweder landet man bei Begründungsversuchen in einem infiniten Regreß oder in einem Zirkel oder man bricht die Begründungen dogmatisch ab. Und für jeden Zweig des Trilemmas scheint sich zu ergeben, daß überhaupt keine Begründung zustande kommt. Wie schon erwähnt, ergreift keiner die Option des infiniten Regresses. Fundamentalisten verfolgen die dogmatische Position mit der Behauptung, es gebe so etwas wie selbstbegründende Aussagen, an denen der Regreß legitimerweise abgeschnitten werden darf; das könnten entweder allgemeine Prinzipien wie logische oder apriorische Wahrheiten sein oder aber Beobachtungsaussagen wie etwa „ich sehe jetzt auf ein bedrucktes Papier". Kohärentisten wollen hingegen die Zirkularität als erträgliche oder plausible Alternative aufbauen. Die fünfte Bedingung liegt jedenfalls in der Frage, wie gut ein Begründungsbegriff mit diesem Trilemma zu Rande kommt.[80]

Um die fünf Gedanken, die für den Begriff rationaler Begründung wesentlich sind, nochmals festzuhalten:

1. Gründe führen mit relativer Zuverlässigkeit zu wahren Ergebnissen.
2. Gründe sind bewertbar und werden vom Subjekt stets als gute Gründe bewertet.
3. Es gibt starke und schwache Gründe und auch Gegengründe.
4. Es ist zwischen „*A* ist Grund für *B*" und „*S* hat *A* als Grund für *B*" zu unterscheiden.
5. Begründungen sind mit dem Begründungstrilemma konfrontiert.

Im weiteren will ich nun die vier in der philosophischen Tradition hauptsächlich vorfindlichen Grundtypen an Begründungsbegriffen darstellen – wie gesagt, nur die Grundtypen; im Detail gibt es viele Mischformen und subtile Ausdifferenzierungen – und jeweils mit der eben erstellten Liste abgleichen, um auf diese Weise die Unterschiede und die Vor- und Nachteile dieser verschiedenen Begriffe klarer in den Blick zu bekommen. Bei diesen Typen handelt es sich um

den deduktiven Begründungsbegriff,
den kausalen Begründungsbegriff,
den kohärentistischen Begründungsbegriff und
den, wie ich ihn nenne, induktiven Begründungsbegriff.

[80] BonJour (1985) befaßt sich ausschließlich und in sehr gründlicher Weise mit dem Begründungstrilemma und argumentiert darin für eine modifizierte kohärentistische Position.

Der deduktive Begründungsbegriff, der wesentlich auf den Begriff der logischen Folgerung Bezug nimmt, ist vielleicht der nächstliegende. Er läßt sich so erklären:

> *A* ist ein Grund für *B* genau dann, wenn *B* logisch aus *A* folgt (und nicht vice versa).

Hierzu ist zunächst festzuhalten, daß die so erklärte Relation deduktiver Begründung irreflexiv, asymmetrisch und transitiv ist.[81] Wie verhält sich nun die deduktive Begründung zu unserer Liste von Bedingungen?

Als erstes fällt auf, daß die dritte Bedingung, die Stärkegrade von Gründen anerkennt, auf den deduktiven Begründungsbegriff nicht paßt. Es gibt keine Grade logischer Folgerung; und damit geht diese Differenzierung verloren. Zudem könnte man zwar deduktive Gegengründe erklären: *A* ist Gegengrund für *B* gdw. non–*B* aus *A* logisch folgt. Doch weil es keine Stärkegrade gibt, taugt auch dieser Begriff nichts; denn ohne sich einer Inkonsistenz schuldig zu machen, kann man danach nicht zugleich über Gründe und Gegengründe für eine Proposition verfügen.

Die zweite Bedingung ist freilich erfüllt; zwingendere als deduktive Gründe kann es gar nicht geben. Ebenso ist die erste Bedingung erfüllt: deduktive Begründungsschritte sind maximal zuverlässig, da wahrheitskonservierend; mit dem Gründen ist auch das Begründete notwendig wahr. Die Kehrseite davon ist, daß sich damit mögliche Zweifel am Begründeten automatisch auch auf den Grund dafür übertragen. Auch die vierte Bedingung ist erfüllt; die deduktive Begründungsbeziehung besteht unabhängig davon, ob Grund oder Begründetes vom Subjekt akzeptiert werden.

Was schließlich das Begründungstrilemma angeht, den fünften Aspekt, so ist ein Verfechter des deduktiven Begründungsbegriffes aufgrund seiner formalen Eigenschaften ersichtlich auf die Position des Dogmatismus festgelegt. Und hier bekommt er ein großes Problem. Was könnten die Basisüberzeugungen sein, die für sich evident sind und bei denen jedenfalls die Begründung abgebrochen werden darf? Hätte man welche, würde sich diese Evidenz auf alles daraus Begründete übertragen. Das scheint aber nicht generell die epistemische Situation zu sein, in der wir uns be-

---

[81] Eine Relation R ist irreflexiv, wenn für alle x aus dem Bereich der Relation $\neg xRx$ gilt; eine Relation R ist asymmetrisch, wenn für alle x,y aus dem Bereich der Relation gilt: wenn xRy, so $\neg yRx$; und eine Relation ist transitiv, wenn für alle x,y,z aus dem Bereich der Relation gilt: wenn xRy und yRz, so xRz.

finden. Bei unseren empirischen Überzeugungen beschränken wir uns jedenfalls nicht auf deduktive Folgerungen aus unseren Wahrnehmungsüberzeugungen; und nach dem deduktiven Begründungsbegriff sind eben, wie gerade festgestellt, die Gründe eben nicht sicherer als das Begründete.

Allenfalls in der Mathematik finden wir uns in dieser epistemischen Lage; da scheint man von Sicherem zu Sicherem deduktiv fortschreiten zu können, wenn man den Beweis eines Theorems auf der Basis von Axiomen eingesehen hat. Doch selbst im Falle der Mathematik sind unsere Evidenzgefühle alles andere als zuverlässig. Generationen von Philosophen galt die euklidische Geometrie als selbstevidentes Axiomensystem. Doch haben die Mathematiker im 19. Jahrhundert Alternativen entwickelt, und die Physiker erzählen uns mittlerweile, daß diese Geometrie, in großem Maßstab betrachtet, nicht einmal eine empirisch korrekte Darstellung der Welt gibt. Zudem geschah es immer wieder, daß Mathematiker ein Axiomensystem präsentierten, das sich dann nach langer kritischer Untersuchung als widersprüchlich erwies – es sei etwa nur an Freges Versuch, die Mathematik auf die Logik zurückzuführen, erinnert.

All das läßt als Diagnose angemessen erscheinen, daß der deduktive Begründungsbegriff nur einen Teil unseres Begründungsvokabulars abdeckt. Seine Vorteile sind unübersehbar, aber er hat auch seine Beschränkungen, wie sich gerade an der epistemologisch so wichtigen fünften Bedingung zeigt. Generell habe ich den Eindruck, daß der deduktive Begründungsbegriff gerade in der Philosophie des 20. Jahrhunderts eher eine zu große Aufmerksamkeit gefunden hat – was aber andererseits angesichts der ungeheuren Entwicklung der Logik, hinter der andere erkenntnistheoretische relevante Gebiete arg hinterherhinkten, nicht verwunderlich ist. Der Schwerpunkt des Interesses sollte sich aber, wie ich plausibel zu machen hoffe, verlagern.

Als nächster sei der kausale Begründungsbegriff betrachtet. Wenn man fragt, warum ein Subjekt *S* die Überzeugung, daß *p*, hat, so kann man das auch als Frage danach verstehen, was die Überzeugung, daß *p*, in *S* hervorrief. Die Antwort auf diese Frage scheint jedenfalls dann, wenn zu dieser Antwort weitere Überzeugungen von *S* angeführt werden, Gründe zu liefern, die *S* für seine Überzeugung, daß *p*, hat. Dieser Gedanke stand übrigens auch schon hinter der externalistischen Wissenstheorie Alvin Goldmans, die wir im Kapitel 3 behandelt haben. Genauer ist der Begriff so formuliert:

> *A* ist ein Grund für *B* gdw. die Tatsache, daß *S* die Überzeugung, daß *A*, hat, eine Ursache dafür ist, daß *S* die Überzeugung, daß *B*, hat.

Das bedeutet, daß wir in unserem Überzeugungssystem Ursachen und Wirkungen voneinander unterscheiden müssen. Diese Unterscheidung ist intuitiv naheliegend. Wie aber läßt sie sich in den in den letzten zwei Kapiteln diskutierten Repräsentationen doxastischer Zustände unterbringen? Am besten dadurch, daß man die dargestellte Dynamik kausal deutet: die akzeptierte Information ist die Ursache, und die dadurch gemäß den Konditionalisierungsregeln erzeugten neuen oder veränderten Überzeugungen sind die Wirkungen. Doch scheint das noch zu grob; eine Information erzeugt danach viele Änderungen in einem einzigen Kausalschritt, während man intuitiv eher erwartet, daß sich die Änderungen in mehr oder weniger langen Kausalketten durch das Überzeugungssystem fortpflanzen. Mit der Theorie Bayes'scher Netze läßt sich dieser Intuition eher gerecht werden; aber es ist doch darauf hinzuweisen, daß hier ein Problem liegt, das meines Wissens auch in der Literatur nicht gut behandelt worden ist.

Um zunächst wieder die formalen Eigenschaften festzuhalten: Die kausale Begründungsrelation ist wie die deduktive irreflexiv, asymmetrisch und transitiv, einfach weil die Kausalbeziehung diese Eigenschaften aufweist. Wie verhält sie sich angesichts unserer fünf Bedingungen?

Der vierten Bedingung, der Unterscheidung zwischen Grund-Sein und Grund-Haben, ist nicht Genüge getan, weil sich der kausale Begründungsbegriff eben nur auf tatsächliche Überzeugungen bezieht; für Propositionen, die nicht Inhalt tatsächlicher Überzeugungen sind, ist er nicht erklärt. Dieses Manko könnte man jedoch vielleicht damit beseitigen, daß man sagt, daß *A*, sofern *S* es nicht akzeptiert, dann ein Grund für *B* ist, wenn die Akzeptation von *A* in *S* die von *B* verursachen *würde*.

Auch die Unterscheidung zwischen starken und schwachen Gründen droht, jedenfalls unter einem deterministischen Bild der Kausalbeziehung, zu verschwinden; unter dieser Voraussetzung ähnelt die Situation sehr der Lage beim deduktiven Begründungsbegriff. Geht man jedoch von einer probabilistischen Konzeption von Kausalität aus, so kann man stärkere und schwächere Ursachen und auch Gegenursachen unterscheiden und damit auch von schwächeren und stärkeren Gründen und von Gegengründen reden.

Was das Begründungstrilemma angeht, so ergibt sich auch unter dem kausa len Begründungsbegriff die dogmatische Auflösung, weil es keine kausalen Zirkel geben kann. Die Basis, bei der man dann landet, besteht offenbar gerade aus all unseren Wahrnehmungsüberzeugungen oder -urteilen. In die Diskussion, ob diese in der Tat als erkenntnistheoretische Basis taugen, will ich mich nun nicht ernsthaft verstricken.[82] Doch ist viel eher als beim deduktiven Begründungsbegriff zu hoffen, daß man hier auf eine Basis stößt, die einerseits begründungsstiftend und andererseits aber nicht weiter begründungspflichtig ist.

Wie steht es dann gemäß diesem kausalen Begriff mit der ersten Bedingung, d.h. der Zuverlässigkeit von Begründungen? Das ist nicht so klar; die Wahrheitsträchtigkeit dieser mentalen Kausalprozesse der Meinungsbildung ist vorderhand ganz offen. Doch ist in dem naturalistischen Rahmen, in den man sich mit diesem Begründungsbegriff begibt[83], das folgende evolutionäre Argument beliebt und plausibel: Arten von Lebewesen, deren kognitive Prozesse nicht im wesentlichen zuverlässig sind, hätten den Millionen von Jahren währenden Überlebenskampf nicht bestanden. Teil der Erklärung des offenkundig (am Ende womöglich uns selbst?) überwältigenden Erfolges der menschlichen Spezies muß also sein, daß unsere Meinungsbildungsprozesse im wesentlich zuverlässig sind. Zum Beispiel verwendet Nozick (1993), Kap. IV, dieses Argument; Stich (1990), Kap. 3, unterzieht es einer kritischen Prüfung.

Wie gut dieses Argument ist, will ich hier dahingestellt sein lassen; denn die meines Erachtens entscheidende Schwachstelle des kausalen Begründungsbegriffs zeigt sich an der noch nicht diskutierten zweiten Bedingung unserer Liste: der Bewertungsaspekt löst sich bei ihm schlichtweg auf. Daß das Haben einer Überzeugung das Vorliegen anderer Überzeugungen verursacht, sagt überhaupt nichts darüber aus, ob erstere einen guten Grund liefert; was unsere Gründe sind, wird so zu einer rein empirischen Frage. Das Problem zeigt sich noch auf andere Weise: Unsere Überzeugungen werden ja auf vielfältige Weise verursacht, durch Vergeßlichkeit, durch Wunschdenken, durch Alkohol, etc. Nach dem kausalen Begründungsbegriff sind diese Ursachen keine Gründe, da nur ursächliche Überzeugungen auch Gründe sind.

---

[82] BonJour (1985) kommt letztendlich zu einer negativen Schlußfolgerung und versucht dann der grundlegenden erkenntnistheoretischen Rolle von Wahrnehmungen, die man ja nicht leugnen kann, auf kohärentistische Weise gerecht zu werden.

[83] Von naturalisierter Erkenntnistheorie spricht man dann, wenn die Erkenntnistheorie als eine rein empirische Disziplin ohne besonderen Status neben den anderen empirischen Wissenschaften betrachtet wird, die eben Meinungs- und Erkenntnisbildung als einen natürlichen empirische Prozeß studiert.

Aber sicherlich können auch Überzeugungen sozusagen auf krummem Wege neue Überzeugungen hervorrufen, und erstere wären dann intuitiv klarerweise keine Gründe für letztere. Also muß der kausale Begriff noch zwischen begründungsstiftenden und nicht begründungsstiftenden Kausalbeziehungen differenzieren. Diese Differenzierung ist schwierig und kann meines Erachtens die evaluative Frage nach den guten Gründen letztlich nicht umgehen.[84]

Der nächste Kandidat ist der kohärentistische Begründungsbegriff. Von allen hier dargestellten Begriffen ist er der am wenigsten präzise, wie man sogleich an der Definition sieht:

> *B* ist begründet gdw. *B* zum restlichen Überzeugungssystem paßt, d.h. mit ihm kohärent ist.

Hier haben wir nicht einmal mehr eine Relation des Begründens vorliegen. Die Leitidee des Ansatzes besteht darin, daß Begründung eine holistische Angelegenheit ist; im Prinzip können und müssen alle Überzeugungen zur Begründung einer weiteren herangezogen werden; und so ist die Stützung letztendlich immer wechselseitig. Ein Beispiel: ich bin im Urlaub in der Toskana, sitze in einer Trattoria und wundere mich gar sehr, als am andern Ende des Raumes ein Mensch erscheint, der aussieht wie mein Freund, spricht wie mein Freund etc.. Nur: vorgestern saß ich mit ihm noch in München zusammen und er erzählte, tags darauf zum Urlaub an die Nordsee zu fahren. Ist dieser Mensch da nun mein Freund oder ist er es nicht? Nun kann ich meine Überzeugungen überprüfen: er sieht so aus, spricht so – das ist Beleg dafür, daß er es ist. Aber warum erzählt er mir dann, er würde an die Nordsee fahren? Ich kenne ihn als absolut vertrauenswürdig und aufrichtig – das kann nicht sein, daß er jetzt auf einmal in der Toskana weilt. Und die Frau an seiner Seite hier, das ist nicht seine Freundin, mit der er seit Jahren zusammen ist und alles gemeinsam tut. Nein, dieser Gast ist der Doppelgänger meines Freundes, usw.. Ein anderes Beispiel: viele werden das Gefühl kennen, aus einem sehr realistischen Traum zu erwachen und nicht zu wissen, ob es ein Traum war oder nicht. Der Test besteht dann in einer Überprüfung dessen, ob der Trauminhalt zum restlichen Überzeugungssystem paßt. (Einen Skeptiker berührt das alles wenig; angesichts der im 2. Kapitel dargestellten Überlegungen von Descartes und Hume wird er darauf verweisen, daß Kohärenz allein nicht den Wahrheitsgehalt des Überzeugungssystems garantiert. Allerdings geht es momentan

[84] Vgl. dazu auch Spohn (1993).

nicht um den Abweis skeptizistischer Überlegungen; natürlich kann eine Überzeugung sehr gut begründet und dennoch falsch sein.)

Die Überprüfung unserer Kriterienliste steht diesmal unter dem Vorzeichen, daß der kohärentistische Begriff gar keine Relation zum Ausdruck bringt. Wenn eine Überzeugung zum restlichen System paßt, wird sie damit auch als gut akzeptiert sein; die Rede vom Passen ist ebenso evaluativ wie die von guten Gründen. Der zweiten Bedingung ist also Genüge getan. Der dritte Gesichtspunkt hinsichtlich der starken und schwachen und der Gegengründe ist auch gewährleistet; zu manchen Überzeugungssystemen paßt eine Überzeugung, zu anderen nicht, und zu manchen besser, zu manchen schlechter. Der vierte Punkt, die Unterscheidung von „Grund sein" und „einen Grund haben" läuft zunächst leer, einfach in Ermangelung eines Grundes als Relationsglied; doch könnte man als Ersatz dafür die Unterscheidung zwischen Propositionen, die in einem Überzeugungssystem nicht enthalten sind, aber dazu passen *würden*, und Überzeugungen, die zu einem System gehören und da auch hineinpassen, anbieten.

Es liegt nahe, daß sich der kohärentistische Begründungsbegriff angesichts des Begründungstrilemmas kohärentistisch verhält; irgendwie soll ja jede Überzeugung durch die jeweils anderen insgesamt gestützt sein. Um das genauer verstehen zu können, müßte allerdings die Begrifflichkeit klarer sein. Die Unklarheit macht sich auch bemerkbar, wenn wir zuletzt die erste Bedingung betrachten; solange nicht erklärt ist, was „passen" heißen soll, kann man die Zuverlässigkeit des Verfahrens gar nicht beurteilen. Was also kann „passen" hier bedeuten?

Der erste Vorschlag ist, daß es soviel bedeutet wie „logisch folgen". Aber daraus resultiert ist ein zu enger Kohärenzbegriff. Bei der Geschichte mit dem Freund in der toskanischen Trattoria stelle ich fortgesetzt Betrachtungen an, die gar keine logischen Folgerungen sind. Meine Meinung, er sei absolut aufrichtig, begründet zwar meine Ansicht, daß er dann tut, was er sagt; aber dies folgt nicht logisch. Also betrachten wir einen zweiten Vorschlag: „passen" heißt soviel wie „konsistent sein mit dem restlichen System". Das ist jedoch zu weit; sowohl die Überzeugung, mein Freund sei in der Toskana als auch die, es sei bloß sein Doppelgänger, sind letztlich konsistent mit meinen restlichen Überzeugungen.

Die beiden bisherigen Vorschläge markieren die präzisen, aber untauglichen Extrempunkte im Spektrum möglicher Explikationen; was dazwischen liegt, bleibt leider eher schwammig. Manche Philosophen haben versucht, präziser darzulegen, was sie unter Kohärenz verstehen.[85] So verweist Lehrer (1990) darauf, daß so etwas wie explanatorische Kohärenz wesentlich sei; je mehr die Elemente eines Überzeugungssystems durch Erklärungen miteinander verknüpft sind, desto mehr gewinnt das System an Kohärenz. Damit verschiebt Lehrer die Last der Explikation auf den Erklärungsbegriff; doch ist diese schwierig und uneindeutig genug, und ob das Material, das im Anschluß an die Untersuchungen von Hempel dazu aufgehäuft wurde, diese Last tragen kann, ist nicht klar.[86] Der unklare Kohärenzbegriff bleibt also die zentrale Schwäche des kohärentistischen Begründungsbegriffs.

Nun gilt es noch den letzten der Begründungsbegriffe zu untersuchen, den ich oben als den induktiven bezeichnet habe, auch wenn das vielleicht nicht die Standardbezeichnung für diesen Begriff ist, der in diesem Jahrhundert neben dem deduktiven Begründungsbegriff eher ein stiefmütterliches Dasein fristete. Für ihn gibt es wieder eine relationale Explikation:

A ist ein Grund für B gdw. $A \neq B$ und A positiv relevant für B ist.

Was bedeutet hier positive Relevanz? Umgangssprachlich heißt das, daß das Begründete mit dem Grund eher oder stärker geglaubt wird als ohne den Grund. In einem probabilistischen Ansatz (wobei P ein subjektives Wahrscheinlichkeitsmaß sei) läßt sich das so präzisieren: A ist positiv relevant für B (relativ zu P), wenn $P(B|A) > P(B)$ bzw. $P(B|A) > P(B|\neg A)$; wie leicht nacuhzuprüfen ist, sind die beiden Formulierungen äquivalent, wenn $P(A) > 0$ und $P(\neg A) > 0$. Auch im Rahmen qualitativer Modelle läßt sich ein Begriff positiver Relevanz spezifizieren, sofern diese Modelle in der Lage sind, bedingte Überzeugungen und Überzeugungsänderungen zu erfassen. Im Rahmen der Theorie der Rangfunktionen sieht das so aus: A ist ein Grund für B relativ zur Rangfunktion κ genau dann, wenn $\kappa(\text{non-}B|A) > \kappa(\text{non-}B|\text{non-}A)$ oder $\kappa(B|A) < \kappa(B|\text{non-}A)$. Oder um es mit Hilfe der κ zugeordneten Glaubensfunktion β durchsichtiger auszudrücken: A ist ein Grund für B relativ zu κ gdw. $\beta(B|A) > \beta(B|\text{non-}A)$.[87]

---

[85] Cf. Rescher (1973), der den Kohärenzbegriff im Kontext der Wahrheitstheorie untersucht hat, oder die schon zitierten Bücher von Bonjour und Lehrer. Für Lehrers Hinweise auf explanatorische Kohärenz cf. besonders S. 95.

[86] Für einen gründlichen aktuellen Überblick über die Erklärungsliteratur vgl. Salmon (1989).

[87] Cf. hierzu Spohn (1988) und (1991). Dieser Begründungsbegriff taugt eher als der in analoger Weise im Rahmen von Gärdenfors' Überzeugungsänderungstheorie eingeführte Begriff, der einige unangemessenen Merkmale aufweist.

Die so charakterisierte induktive Begründung umfaßt die deduktive als Spezialfall: Wenn A ein deduktiver Grund für B ist, heißt das ja, daß B aus A logisch folgt. Dann aber gilt $P(B|A) = 1$. Wenn wir das Regularitätsaxiom fordern, gilt auch, daß $P(B) < 1$, sofern B eine kontingente Proposition ist. Wenn B also kontingent und A ein deduktiver Grund für B ist, dann ist A auch ein induktiver Grund für B. Doch verhält sich die induktive Begründungsrelation relationstheoretisch ganz anders als die deduktive: sie ist irreflexiv, symmetrisch und nicht transitiv. Daß die Relation irreflexiv ist, sieht man sofort; die Symmetrie läßt sich schnell zeigen: nach der Definition der bedingten Wahrscheinlichkeit gilt ja $P(B|A) > P(B)$ gdw. $P(A \wedge B) > P(A) \cdot P(B)$ gdw. $P(A|B) > P(A)$. Nun ist noch die Nicht-Transitivität zu behandeln; sie ist mit einem Gegenbeispiel gezeigt, in dem A ein induktiver Grund für B und B ein induktiver Grund für C, A aber kein induktiver Grund für C ist. Das bedeutet, daß $P(B|A) > P(B)$, $P(C|B) > P(C)$ und $P(C|A) \leq P(C)$ gelten müssen. Diese drei Ungleichungen sind aber ohne weiteres gleichzeitig erfüllbar. Ein Beispiel: Wir haben zehn von 1 bis 10 durchnumerierte Kugeln. Seien A, B, C die Eigenschaften, eine der Nummern aus den jeweiligen Mengen zu haben, wobei $A = \{1,3,5,7,9\}$, $B = \{1,3,5,6,8\}$ und $C = \{2,4,5,6,8\}$. Betrachten wir eine beliebige Kugel, deren Nummer unbekannt ist. Mit welcher Wahrscheinlichkeit hat sie eine der drei Eigenschaften? Wir haben $P(A) = P(B) = P(C) = 0{,}5$. Dann ist, die Eigenschaft A zu haben, positiv relevant dafür, die Eigenschaft B zu haben, denn $P(B|A) = 0{,}6 > P(B)$; die Eigenschaft, B zu haben, ist positiv relevant für das Vorliegen der Eigenschaft C, da $P(C|B) = 0{,}6 > P(B)$. Aber A ist nicht positiv relevant für C, denn $P(C|A) = 0{,}2$, was kleiner ist als $P(C)$. Diese formalen Eigenschaften muß man sich gründlich klar machen; wenn der induktive Begründungsbegriff angemessen ist, so gilt es manche Begründungsgewohnheiten sorgfältig zu überprüfen.

Ich führte schon aus, daß der deduktive vom induktiven Begründungsbegriff umfaßt wird. Das gleiche läßt sich, wenn auch vager, für die anderen Begriffe sagen. Denn es dürfte plausibel sein, daß der induktive Begründungsbegriff, für den ich ja zwei untadelige Präzisierungen angeboten habe, ein guter, vielleicht sogar der relativ beste Kandidat ist, um den kohärentistischen Begründungsbegriff mit seiner vagen Rede vom kohärenten Passen weiter zu explizieren. Und was den kausalen Begründungsbegriff anlangt, so ist auch dieser im induktiven enthalten. Der induktive Begründungsbegriff hängt ja eng mit dem Begriff rationaler Überzeugungsveränderung zusammen; gemäß den gegebenen Definitionen der Konditionalisierung und der positiven zieht der Erwerb von Gründen offensichtlich doxastische Änderungen nach sich. Diesen Änderungsprozeß kann man natürlich kausal deuten. Das heißt: vorausgesetzt, daß die tatsächliche (kausale) Dynamik der rationalen Dynamik von Überzeugungen folgt, ist der kausale Prozeß ein Fall des induktiven Änderungsprozesses. Wichtig ist dabei, daß man diese Voraussetzung nur vom Standpunkt einer Rationalitätstheorie aus machen kann; unter einer rein kausalen naturalistischen Sichtweise ist diese Voraussetzung unzugänglich (was mir ja das entscheidende Manko dieser Sichtweise zu sein schien). Daß der induktive Begründungsbegriff die anderen auf diese Weise in sich aufnehmen kann, ist schon ein gutes Zeichen.

Wie verhält sich dieser Begriff nun zu den aufgelisteten Bedingungen? Die Punkte 2 und 3 sind erfüllt: induktive Gründe lassen sich so verstehen, daß sie vom Subjekt jeweils als gute Gründe bewertet sind; und die Unterscheidung von starken und schwachen sowie von Gegengründen ist ersichtlich leicht zu vollziehen.

Hinsichtlich der vierten Bedingung muß man etwas aufpassen. Was heißt es, daß einer Grund A für B *hat*? Offenbar, daß er von A überzeugt ist. Im probabilistischen Rahmen könnte man das dadurch ausdrücken, daß $P(A) = 1$. Relativ zu einem solchen P kann aber A gemäß obiger Erklärung kein Grund für B sein, da dann $P(B|A) = P(B)$ und $P(B|\neg A)$ gar nicht definiert ist. Ein Ausweg daraus ist, daß man auf dem Regularitätsaxiom besteht und sich weigert, kontingenten Propositionen die extremen Wahrscheinlichkeiten 0 und 1 zuzuweisen; dann wären die einschlägigen Wahrscheinlichkeiten, die positive Relevanz charakterisieren, immer definiert. Allerdings müßte man dann auch sagen, daß man den Grund A schon hat, wenn $P(A) = 1-\varepsilon$ für ein geeignet gewähltes kleines ε – womit man in gefährliche Nähe des Lotterieparadoxes gerät. Insofern ist es vielleicht der bessere Ausweg, wiederum zu den früher angedeuteten Popper-Maßen Zuflucht zu nehmen. Im rangfunktionalen Rahmen ergeben sich freilich keine derartigen Probleme.

Offenbar ist die induktive Begründungsrelation auf eine kohärentistische Auflösung des Begründungstrilemmas verpflichtet, wie man schon an der Symmetrie dieser Relation erkennt. Wie bei den anderen Begriffen läßt sich daraus aber auch hier keine Bewertung ableiten. Eine solche Bewertung könnte sich nur aus einer gründlichen Diskussion (etwa im Format von BonJour 1985) dessen ergeben, wie man sich überhaupt fundamentalistische oder kohärentistische Lösungen des Trilemmas verständlich machen kann. Diese Diskussion habe ich nur geleistet; so weit habe ich nur deutlich gemacht, auf welche Option man sich mit diesem oder jenen Begründungsbegriff konzentrieren muß.

Es bleibt als letztes die Frage nach der ersten Bedingung, nach der Verläßlichkeit induktiver Gründe. Diese scheint mir vorderhand nicht sichergestellt. Eine Methode wäre, sie durch zusätzliche Rationalitätsaxiome zu gewährleisten; das im letzten Kapitel Reichenbach-Axiom wäre etwa ein solches Axiom. Doch wäre auch diese Frage eigentlich viel gründlicher unter die Lupe zu nehmen.

So hat dieses Kapitel zwar die Neigung des Autors erkennen lassen, welche Route er durch das Gestrüpp der Begründungsthematik für die richtige hält; und es hat auch einige Gründe (!) für diese Neigung genannt. Das eine feste Schlußfolgerung etabliert worden wäre, kann man aber beim besten Willen nicht behaupten. So war mein Anliegen hier eher, die Begründungsthematik mit Hilfe zweier Raster aufzufächern, mit den früheren Kapiteln zu verknüpfen und so weiterem und vielleicht informierterem oder fruchtbarerem Nachdenken anheim zu geben.