Bayesianisches Schliessverfahren im Kontext von Intelligent Design

Christian Knobel

28. November 2011

Inhaltsverzeichnis

1. Einleitung

Es ist nun über zehn Jahre her, als die „Intelligent Design“ (ID) Bewegung in Amerika mit dem Buch von Michael Behe „Darwin's black box“ einen dramatischen und medienwirksamen Aufschwung erlebt hat. Seither wird ID in Amerika, aber auch in Europa äusserst kontrovers diskutiert. Dieser Essay ist als Diskussionsbeitrag gedacht, der dafür plädiert, ID und Ursprungsfragen generell aus der Perspektive des Bayesianischen Schliessverfahrens (BS) zu betrachten, um die Plausibilität gewisser Theorien oder Paradigmen gegen andere abzuwägen.

Das BS ist im Prinzip eine Rechenvorschrift, wie die Plausibilitäten von Aussagen (z.B. von Theorien oder Hypothesen) von den Plausibilitäten anderer Aussagen abgeleitet werden können. Insbesondere sagt das BS wie die Plausibilität einer bestehenden Aussage verändert wird, wenn neue Evidenz (oder Beweismaterial) in die Betrachtung einbezogen wird. Es handelt sich damit um eine Hilfe für das korrekte Argumentieren mit Aussagen, wenn diese mit Unsicherheiten behaftet sind, wie dies im Spannungsfeld der Ursprungsfragen häufig der Fall ist. Das BS ist aber keinesfalls auf die Anwendung auf Ursprungsfragen eingeschränkt, sondern kann universell eingesetzt werden.

Das besondere Merkmal des BS liegt in seiner präzisen Struktur und in der Tatsache, dass es auf beliebige Fragestellungen anwendbar ist, d.h. auch auf solche, die über die Naturwissenschaft hinausgehen. Das BS ist insbesondere ein „logisch korrektes“ Schliessverfahren, da es in Übereinstimmung mit den Axiomen der Wahrscheinlichkeitstheorie steht. Diese Eigenschaft impliziert auch, dass es anderen Schliessverfahren, die nicht logisch zwingend sind, wie beispielsweise der Abduktion oder dem Analogieschluss auf einen Designer, überlegen ist. Ich möchte mit diesem Essay zu einer im Spannungsfeld der Ursprungsfragen logisch korrekten Argumentationsweise ermutigen und deutlich machen, wie weltanschauliche Prämissen die Argumentationsweise einer Person beeinflussen können.

Dieser Essay besteht aus drei Kapiteln. Im ersten längeren Kapitel wird das BS in allgemeiner Form eingeführt, ohne im Speziellen auf die Ursprungsproblematik Bezug zu nehmen. Im zweiten Kapitel, das kürzer ist als das erste, wird dieses Schliessverfahren auf ID angewendet. Es werden dabei keine abschliessenden Resultate erzielt, sondern beispielhaft gezeigt, wie das BS auf die Ursprungsproblematik angewendet werden kann und welche Einsichten sich damit gewinnen lassen. Im dritten Kapitel werden schliesslich die Haupterkenntnisse zusammengefasst.

Zusätzliches (teilweise technisches) Material, das für die Hauptaussagen dieses Essays nicht notwendig ist, wurde in die Anhänge gestellt. In einigen Fussnoten finden sich zudem teilweise technische Anmerkungen, die für das Verständnis des Haupttextes aber nicht notwendig sind. Sie dürfen vom Leser getrost übergangen werden. Da die Hauptkonzepte dieses Essays auf ein paar Grundprinzipien aufgebaut werden können, ohne dass auf Ausführungen in der Literatur zurückgegriffen werden muss, wurde die angeführte Literatur auf ein Minimum beschränkt.

2. Das Bayesianische Schliessverfahren (BS)

Über Bayesianismus wurden unzählige Bücher und Artikel geschrieben. Über einzelne Aspekte gibt es seit Jahrhunderten Kontroversen, die bis heute nicht abgeschlossen sind. Dieser Aufsatz verfolgt nicht das Ziel, die verschiedenen Standpunkte dieser Kontroversen anhand der Literatur aufzuarbeiten, sondern eine Einführung in die wichtigsten Konzepte zu geben mit dem

Schwerpunkt, diese Konzepte auf die Ursprungsfrage anzuwenden. Es wird dafür vor allem der Darlegung in den Kapiteln drei und vier von Richard Swinburne (2001) „Epistemic Justification" gefolgt, welche die Grundlagen des BS sorgfältig und umfassend behandelt und auch viele kontroverse Punkte diskutiert. Ich stimme zwar nicht notwendigerweise mit jedem Detail von Swinburnes „Epistemic Justification" überein, dennoch sei dieses Werk interessierten Lesern als vertiefende Lektüre wärmstens empfohlen.[1]

Im Abschnitt 2.1 werden „Aussagen" als die grundlegenden Bestandteile eingeführt, mit welchen das BS arbeitet.[2] Im Abschnitt 2.2 wird die Interpretation der mit dem BS verbundenen „Wahrscheinlichkeit" diskutiert und im Abschnitt 2.3 wird das BS mittels Axiomen definiert, wobei als Argumentationshilfe Bayes Theorem hergeleitet wird. Im Abschnitt 2.4 wird schliesslich noch auf den Problempunkt der Prior-Wahrscheinlichkeit eingegangen und im Abschnitt 2.5 führen wir zur Verständnishilfe ein Rechenbeispiel ausführlich vor.

2.1. Aussagen

Beim BS geht es darum, beliebige Aussagen über „Gott und die Welt" auf ihre Plausibilität zu überprüfen, indem ihnen „Wahrscheinlichkeiten" zugeordnet werden. Dabei werden alle (logisch denkbaren) Aussagen zugelassen. Beispiele sind „Die Menschen stammen von affenartigen Vorfahren ab.", „Heute Abend regnet es." oder „Gott ist allmächtig." Grundsätzlich können sogar Aussagen zugelassen werden, die in sich widersprüchlich sind wie z.B. "Heute Abend regnet es und es regnet heute Abend nicht.", aber solche Aussagen werden per Definition die Wahrscheinlichkeit Null aufweisen, d.h. sie können nicht wahr sein. Dagegen können alle Aussagen, die in sich nicht widersprüchlich sind, prinzipiell wahr sein, auch wenn sie uns noch so absurd erscheinen. Beispiele sind „Es gibt fliegende rosarote Elefanten", „Auf dem Mars leben kleine grüne Männchen" oder „Gott hat einen langen Bart". Solchen Aussagen würden jedoch vermutlich (selbst nach eingehender Untersuchung) eine verschwindend geringe Wahrscheinlichkeit zugewiesen, so dass sie im Grunde genommen vernachlässigt werden können. Aber prinzipiell sind solche Aussagen zugelassen und haben eine Wahrscheinlichkeit ungleich Null, solange nicht gezeigt werden kann, dass sie aus irgendeinem Grund in sich widersprüchlich sind.

Es ist möglich, verschiedene Aussagen zu verknüpfen. Wenn p und q beliebige Aussagen bezeichnen, dann sind die verknüpften Aussagen '$r \vee q$' und '$r \wedge q$' wiederum Aussagen, wobei '$\vee$' dem logischen 'oder' ('das eine oder das andere oder beides') und '$\wedge$' dem logischen 'und' ('das eine und zugleich das andere') entsprechen. Sei beispielsweise $q =$ „Evolution hat stattgefunden" und $r =$ „Gott hat die Lebewesen geschaffen", so gilt $q \vee r =$ „Evolution hat stattgefunden oder aber Gott hat die Lebewesen geschaffen oder beides trifft zugleich zu." $q \wedge r =$ „Evolution hat stattgefunden und zugleich hat Gott die Lebewesen geschaffen." Mithilfe dieser Verknüpfun-

[1] In Chalmers (1999, Kapitel 12) wird das BS aus dem Blickwinkel der Wissenschaftstheorie betrachtet. Es wird in jenem Kapitel z.B. gezeigt, wie das BS die Wissenschaftstheorie von Lakatos verständlich macht, wie das BS eine saubere Bewertung von „Ad hoc"-Hypothesen erlaubt und wie sich damit Beispiele aus der Geschichte der Wissenschaft verstehen lassen. Auf der anderen Seite werden auch die Schwachpunkte und Grenzen des BS behandelt. In diesem Essay verwenden wir das BS aber umfassender als im Rahmen der Wissenschaftstheorie und schränken es nicht nur auf naturwissenschaftliche Hypothesen ein.

[2] Für den mathematisch geschulten Leser kann es hilfreich sind, Aussagen als Mengen aufzufassen. Auf diese Weise kann das BS auf der mathematische Mengenlehre aufgebaut werden, wobei es sich im Wesentlichen zur mathematischen Wahrscheinlichkeitstheorie reduziert. Im Appendix A finden sich entsprechende Ausführungen. Wir werden die Diskussion des BS jedoch durchweg mithilfe von Aussagen formulieren.

gen lassen sich nun beliebige Aussagen aus einfachen Aussagebausteinen zusammensetzen oder komplizierte Aussagen in einfachere zerlegen.

Es sei an dieser Stelle darauf hingewiesen, dass Aussagen aufgrund des Kontextes und der Eigenart der menschlichen Sprache oft keine präzisen bzw. vollständigen Informationen geben. Das bedeutet, solche Aussagen können missverständlich sein.[3] Damit eine Aussage wirklich einen wohldefinierten Inhalt hat, sollte sichergestellt werden, dass die Bedeutung der Begriffe unmissverständlich ist oder dass zumindest alle Menschen möglichst dasselbe darunter verstehen. In den oben verwendeten Aussagen q und r müsste man Beispielsweise genau darlegen, was man unter „Evolution“ oder unter „Gott“ versteht. Was bedeutet es zudem, dass „Gott“ etwas „schafft“? Es ist gut möglich, dass gewisse Leute die Aussagen p und q aufgrund ihres unterschiedlichen Hintergrundwissens unterschiedlich verstehen oder gar als nicht vereinbar ansehen. Doch während diese Aussagen für die meisten Leute einen verständlichen Inhalt haben dürften (wenn auch nicht immer derselbe), so dürfte die Aussage „Die Schwurbel, Flumsen und babigen Schwalme haben die Lebewesen geschaffen“ für die allermeisten Menschen völlig inhaltslos sein. In diesem Essay wird daher immer davon ausgegangen, dass eine Aussage eine eindeutige, unmissverständliche Bedeutung hat. In der Praxis ist diese Voraussetzung natürlich nicht immer erfüllt oder möglicherweise gar nicht erfüllbar.

2.2. Der Wahrscheinlichkeitsbegriff

Der Begriff „Wahrscheinlichkeit“ wird in der Umgangssprache und in der (mathematischen und philosophischen) Literatur nicht einheitlich verwendet. Zuerst wird die statistische Wahrscheinlichkeit eingeführt, um sie vom Bayesianischen Wahrscheinlichkeitsbegriff abzugrenzen.

Statistische Wahrscheinlichkeit

Eine häufige Verwendung des Wahrscheinlichkeitsbegriffs ist die „statistische Wahrscheinlichkeit“[4] (in der Datenanalyse häufig als „frequentist approach“ bezeichnet). Dies ist der Wahrscheinlichkeitsbegriff, den man gewöhnlich in der Schule lernt, wenn zum ersten Mal Wahrscheinlichkeit und Statistik behandelt wird. Er lautet wie folgt: Man habe einen wiederholbaren, stochastischen[5] Prozess, der grundsätzlich N verschiedene Ausgänge $A_1, A_2, \ldots, A_N$ haben

[3] Rautenberg (2008) beispielsweise drückt das in der Einleitung seines Lehrbuchs über die mathematische Logik folgendermassen aus:

> Im Alltagsleben hängt die Bedeutung fast jeder Äusserung vom Kontext ab. In den meisten Fällen sind logische Relationen nur angedeutet, selten explizit ausgedrückt. Oft fehlen grundlegende Voraussetzungen [...].

Für die damit verbundenen Probleme sei z.B. auf die sprachphilosophische Einführung im Zusammenhang der Logik in Zoglauer (2008, Kapitel 1) verwiesen.

[4] Swinburne (2001) bezeichnet diese Art von Wahrscheinlichkeit „statistical Probability“ (p. 57). Swinburne ist sehr sorgfältig bei der Verwendung von Begriffen und führt noch weitere Wahrscheinlichkeitsbegriffe ein, die hier aber nicht aufgegriffen werden sollen.

[5] Stochastisch bedeutet hier, dass die nacheinander folgenden Ausgänge des Prozesses nicht miteinander korreliert sind, d.h. der laufende Prozess weiss nichts von den früheren Ergebnissen. Zudem wird der Prozess immer unter denselben, idealistischen Bedingungen wiederholt.

kann. Die Wahrscheinlichkeit p_i, dass der Prozess im Ausgang A_i endet, ist definiert als

$$p_i = \lim_{n\to\infty} \frac{\text{Anzahl Ausgänge} A_i}{\text{Anzahl Wiederholungen} n}\,, \tag{1}$$

wobei n die Anzahl Wiederholungen des Prozesses bezeichnet. In Worte übersetzt bedeutet das einfach, dass im Falle von idealerweise unendlich vielen Wiederholungen der Bruchteil der Ausgänge A_i genau der entsprechenden Wahrscheinlichkeit p_i entspricht. Als beliebtes Beispiel für diese Art von Wahrscheinlichkeit gilt der Wurf einer fairen Münze: Der Wurf einer fairen Münze ist so definiert, dass es nur zwei mögliche Ausgänge gibt – 'Kopf' oder 'Zahl' – wobei beiden genau die Wahrscheinlichkeit 1/2 entspricht. Bei jedem Wurf dieser Münze ist zwar unvorhersehbar, ob 'Kopf' oder 'Zahl' erscheinen wird, wirft man die Münze jedoch immer und immer wieder, so werden etwa die Hälfte der Würfe mit 'Kopf' und die andere Hälfte mit 'Zahl' enden. Diese Art von Wahrscheinlichkeit ist mathematisch wohldefiniert und basiert auf den Axiomen, die von Kolmogorow 1933 postuliert wurden.

Induktive Wahrscheinlichkeit

Für das BS (in der Datenanalyse häufig als „bayesian approach" bezeichnet) wird jedoch ein anderer Wahrscheinlichkeitsbegriff verwendet. Wie im Folgenden deutlich wird, ist er mathematisch nicht mehr im gleichen Sinne wohldefiniert wie die statistische Wahrscheinlichkeit, er ist dafür auf beliebige Aussagen anwendbar und nicht nur auf die Ausgänge eines wiederholbaren stochastischen Prozesses. Seien q und r beliebige Aussagen, so bezeichnet $p(q|r)$ das Ausmass, in welchem die Aussage r die Aussage q bedingt bzw. als wahr erscheinen lässt. Der Ausdruck $p(q|r)$ stellt eine bedingte bzw. konditionierte Wahrscheinlichkeit dar. Oder in anderen Worten: $p(q|r)$ **ist die „Wahrscheinlichkeit", dass die Aussage** q **wahr ist, falls die Aussage** r **als wahr vorausgesetzt wird.** Swinburne bezeichnet diesen Wahrscheinlichkeitsbegriff als „induktive Wahrscheinlichkeit" (Swinburne 2001, p. 62).

Jedem Ausdruck der Form $p(q|r)$ wird nun eine Wahrscheinlichkeit im Intervall $[0, 1]$ zugewiesen. Dabei soll $p(q|r) = 1$ bedeuten, dass die Aussage r die Aussage q logisch impliziert, d.h. wann immer r gilt, so gilt auch q, und $p(q|r) = 0$ soll bedeuten, dass r die Negation von q logisch impliziert, d.h. wann immer r gilt, gilt q sicher nicht. Die Zuweisung von numerischen Werten für die Wahrscheinlichkeit $p(q|r)$ für Fälle, bei denen die Aussage r die Aussage q weder logisch impliziert noch logisch ausschliesst, wird in einem späteren Abschnitt behandelt werden.

Die Interpretation der Wahrscheinlichkeit p ist die „**persönliche Zuversicht**" oder Plausibilität, dass die zugehörige Aussage wahr ist. Diese Zuversicht bzw. Plausibilität zwar auf vornehmlich rationalen Argumenten aufgebaut, aber kann letztlich subjektive Elemente kaum völlig vermeiden.[6] Diese „persönliche Zuversicht" einer Person könnte sich beispielsweise darin äussern, inwiefern die Person bereit ist, für den Wahrheitsgehalt einer Aussage zu wetten bzw. ihre „Hand ins Feuer zu halten". **Je höher der numerische Wert der induktiven Wahrscheinlichkeit, desto höher die persönliche Zuversicht, dass die Aussage wahr ist, und desto höher wiederum die Bereitschaft, auf eine Wette einzusteigen, bei**

[6] Swinburne gehört offenbar zu den Autoren, die hier optimistisch denken. Er räumt zwar ein, dass es kaum möglich sei, $p(q|r)$ für zwei beliebige Ereignisse einen exakten numerischen Wert zuzuweisen, aber er ist der Meinung, dass es beispielsweise wenigstens möglich sei, die Wahrscheinlichkeit relativ zu $p(q'|r')$ für zwei andere Ereignisse q' und r' anzugeben (Swinburne 2001, p. 62).

der man gewinnt, falls sich die Aussage tatsächlich als wahr herausstellen sollte. Natürlich kann in vielen Fällen – und besonders im Kontext der Ursprungsfrage – eine solche Wette nicht durchgeführt werden, da der korrekte Wahrheitsgehalt einer Aussage wohl nie definitiv feststehen wird, aber in diesem Fall kann dennoch die Bereitschaft für einen hypothetischen Wetteinsatz existieren, so dass das Konzept der induktiven Wahrscheinlichkeit sinnvoll ist, wenn auch nicht ganz scharf definiert.

Der induktive Wahrscheinlichkeitsbegriff hat vorerst einmal nichts mit der Häufigkeit von Ausgängen in einem stochastischen Prozess zu tun. Dennoch kann er auf stochastische Prozesse angewandt werden. Man erwartet natürlich, dass in diesem Fall die induktive Wahrscheinlichkeit der statistischen Wahrscheinlichkeit sehr nahe kommt. Idealerweise sind die beiden gleich. Hat man beispielsweise eine faire Münze, so sollte nicht nur die statistische Wahrscheinlichkeit, dass beim nächsten Wurf 'Kopf' erscheint, 1/2 betragen, sondern auch die induktive Wahrscheinlichkeit $p(q|r) = 1/2$, wobei $q =$ „Der nächste Wurf mit dieser Münze ist Kopf" und $r =$ „Diese Münze ist fair" bedeuten. Da es sich hier um ein rein mathematisches Statement handelt, ist es einsichtig, dass die statistische und induktive Wahrscheinlichkeit exakt gleich sind. Das muss aber nicht zwingend so sein. Was ist, wenn ich nicht weiss, ob die Münze fair ist oder nicht? Sagen wir, das einzige, was ich weiss, sind die ersten 100 Würfe einer Serie von Würfen. Aus diesen 100 Würfen geht die Aussage hervor: $r =$ „Es erschienen 61 mal 'Kopf' und 39 mal 'Zahl'." Wie würde ich jetzt meine induktive Wahrscheinlichkeit $p(q|r)$ einschätzen, wobei wiederum gilt $q =$ „Der nächste Wurf mit dieser Münze ist Kopf". Es könnte gut sein, dass die Münze fair ist und dass die Ungleichheit zwischen 'Kopf' und 'Zahl' unter den ersten 100 Würfen eine rein zufällige Schwankung darstellt. Es könnte aber auch sein, dass die Münze leicht gezinkt ist. In diesem Fall ist nicht mehr garantiert, dass meine induktive Wahrscheinlichkeit zwingendermassen $p(q|r) = 1/2$ betragen muss. Also je nach Informationen, die zugrunde liegen, können sich statistische und induktive Wahrscheinlichkeit unterscheiden. Letztlich ist die statistische Wahrscheinlichkeit ein rein idealisiertes, mathematisches Konzept. Man wird nie mit letzter Sicherheit testen können, ob eine Münze wirklich fair ist oder nicht. In der Realität würde letztendlich (ob bewusst oder unbewusst) mit dem Bayesianischen Schliessverfahren die „Zuversicht" angegeben, dass die Münze fair sei.

Induktive Wahrscheinlichkeit in der modernen Physik

Das Konzept der induktiven Wahrscheinlichkeit ist in den Naturwissenschaften schon seit Jahrzehnten etabliert. Es kommt vornehmlich in der Datenanalyse der Kosmologie und Teilchenphysik vor (siehe z.B. Cowan 1998, Kapitel 1.2). In gewissen Gebieten (z.B. bei der Analyse des kosmischen Mikrowellenhintergrundes in der Kosmologie) hat man aus computertechnischen Gründen aufgrund der riesigen Datenmenge fast keine andere Wahl, als auf die Bayesianische Statistik zurückzugreifen, da eine Analyse mittels der „frequentist statistics" viel zu viel Rechenzeit beanspruchen würde. Die Bayesianische Statistik erlaubt es nämlich, numerische Verfahren wie das MCMC (Markov Chain Monte Carlo) anzuwenden, was die Datenanalyse stark beschleunigt. Allerdings ist die Bayesianische Statistik auch heute nicht unumstritten und es gibt beispielsweise in der Kosmologie ganze Konferenzen, die nur dieser Thematik gewidmet sind. Der Grund dafür liegt weniger an der Interpretation des induktiven Wahrscheinlichkeitsbegriffs, sondern vor allem an der Unvermeidlichkeit der sogenannten „prior Wahrscheinlichkeiten", die in einer Analyse letztlich mehr oder weniger apriori gewählt werden müssen. In Appendix B wird dieses Problem weiter vertieft.

2.3. Axiome und Bayes Theorem

Damit der Formalismus des BS auf eine solide Grundlage kommt, sollen nun fünf Axiome eingeführt werden, die zeigen, wie mit Ausdrücken der Form $p(q|r)$ gerechnet werden kann. Seien q, r und s grundsätzlich beliebige Aussagen, so gelten immer die folgenden Axiome (Swinburne 2001, p. 103):

1. $p(q|r) \geq 0$.

   Jeder Wahrscheinlichkeit ist ein Wert grösser oder gleich Null zugeordnet.

2. Falls sich q und r gegenseitig logisch implizieren, dann gilt $p(q|r) = p(r|q)$.

   Falls sich zwei Aussagen q und r gegenseitig logisch implizieren, dann ist die Wahrscheinlichkeit, dass q aus r folgt gleich gross wie dass r aus q folgt.

3. Falls r die Aussage q logisch impliziert, dann gilt $p(q|r) = 1$.

   Falls r die Aussage q logisch impliziert, dann ist die Wahrscheinlichkeit, dass q aus r folgt gleich 1. Das ist der Spezialfall der Deduktion.

4. Ist die Aussage s logisch möglich und sind die Aussagen q und r sich gegenseitig ausschliessend, dann gilt $p(q \vee r|s) = p(q|s) + p(r|s)$.

   Die Wahrscheinlichkeit, dass die Aussagen q oder r wahr sind, falls s als wahr angenommen wird, ist einfach die Summe der Wahrscheinlichkeiten der beiden einzelnen Aussagen p und q ist, falls s als wahr angenommen wird und die beiden Aussagen sich gegenseitig ausschliessen.

5. $p(q \wedge r|s) = p(q|r \wedge s)p(r|s)$.

   Die Wahrscheinlichkeit, dass die Aussagen q und r zugleich wahr sind, falls s als wahr angenommen wird, ist gleich gross wie die Wahrscheinlichkeit, dass nur r wahr ist, falls s als wahr angenommen wird, multipliziert mit der Wahrscheinlichkeit, dass nur q wahr ist, falls diesmal r und s zugleich als wahr angenommen werden.

Die Axiome 1-5 werden manchmal als „Bayesianische Axiome“ bezeichnet und lassen sich intuitiv relativ leicht verstehen (ausser vielleicht das fünfte Axiom). Sie sind aus den Axiomen der herkömmlichen Wahrscheinlichkeitstheorie ableitbar, falls die Aussagen als Mengen im Sinne von Appendix A verstanden werden.[7] Also kann mit der induktiven Wahrscheinlichkeit

[7] Zu Axiom 2 gibt es kein Analog unter den Axiomen der mathematischen Wahrscheinlichkeitstheorie von Kolmogorow, weil zwei Aussagen, die sich gegenseitig logisch implizieren, notwendigerweise die gleichen zugehörigen Mengen aufweisen, d.h. für die Mengen gilt $Q = R$, womit dieses Axiom für Mengen trivialerweise erfüllt ist, d.h. $p(Q|R) = p(Q,Q) = p(R,R) = p(R|Q)$. Axiom 3 reduziert sich für Mengen zudem auf $p(R|R) = 1$. Axiom 5 entspricht in der mathematischen Wahrscheinlichkeitstheorie im Wesentlichen der Definition der bedingten Wahrscheinlichkeit.

letztlich genauso gerechnet werden, wie man es von der mathematischen Wahrscheinlichkeitstheorie gewohnt ist.[8] Es kann von den Axiomen gezeigt werden, dass für zwei sich gegenseitig ausschliessende Aussagen q und r tatsächlich $p(q|r) = p(r|q) = 0$ folgt und für zwei beliebige Aussagen q und r immer gilt $p(q|r) \leq 1$. Damit sind die Axiome mit der Interpretation von p als induktiver Wahrscheinlichkeit verträglich, die wir im letzten Abschnitt eingeführt haben.

Nun wollen wir uns der Frage zuwenden, warum ausgerechnet die oben definierten Axiome 1-5 die Grundlage für ein korrektes Schlussfolgern darstellen sollen. Der Grund ist, dass sie „kohärent" (engl.: coherent, d.h. stimmig) sind, was mithilfe des sogenannten „Dutch Book" Arguments veranschaulicht werden kann. Stellen wir uns ein Pferderennen vor, bei welchem wir für jedes Pferd die Wahrscheinlichkeit festlegen, dass es gewinnt. Die Wetten laufen dabei folgendermassen ab: Setzen wir einen Euro auf den Sieg eines Pferdes mit zugehöriger Wahrscheinlichkeit p, so erhalten wir im Falle eines Sieges zusätzlich $1/p - 1$ des Einsatzes und im Fall einer Niederlage verlieren wir unseren Einsatz. Auf der anderen Seite sei es auch erlaubt, auf die Niederlage eines Pferdes zu setzen. Setzt man 1 Euro auf die Niederlage eines Pferdes mit zugehöriger Wahrscheinlichkeit p, so verliert man im Fall eines Sieges seinen Einsatz und erhält im Falle einer Niederlage $p/(1 - p)$ des Wetteinsatzes. **Dann kann gezeigt werden, dass – ausser man wählt die Wahrscheinlichkeiten für die Pferde in Einklang mit den Axiomen 1-5 – immer jemand so mit dir so wetten könnte (vorausgesetzt du nimmst jede Wette an), dass du mit Sicherheit Geld verlierst, unabhängig welches Pferd tatsächlich gewinnt** (Swinburne 2001, p. 119 ff.). Man sagt, jemand macht ein „Dutch Book" gegen dich.

Ein Beispiel: An einem Pferderennen nehmen drei Pferde teil und jedes Pferd habe die Wahrscheinlichkeit $1/2$ für einen Sieg. Nun könnte ich einen Euro auf den Sieg jedes Pferdes setzen. Am Ende spielt es keine Rolle, welches Pferd gewinnt. Ich verliere zwei Euro, weil zwei Pferde verlieren, und ich gewinne 1 Euro, weil ein Pferd gewinnt. Am Ende habe ich mit Sicherheit einen Euro Verlust gemacht und die Bank hat einen Euro gewonnen. Der Grund dafür liegt darin, dass sich die Wahrscheinlichkeiten der Pferde nicht auf 1 summieren und damit im Widerspruch zu den Axiomen 3 und 4 stehen. In analoger Weise lassen sich Beispiele bilden, die die anderen Axiome betreffen.

Die Tatsache, dass die fünf Axiome kohärent sind im oben beschriebenen Sinne, rechtfertigt ihre Wahl nachträglich. Die Axiome entsprechen der Logik und dem gesunden Menschenverstand, der auch in der mathematischen Wahrscheinlichkeitstheorie zur Geltung kommt. Genauso wie man keine Pferdewette eingehen sollte, wenn die Wettbedingungen nicht den Bayesianischen Axiomen gehorchen, genauso wenig sollte man sein Vertrauen auf die Plausibilität einer Aussage (z.B. einer Theorie oder Hypothese) setzen, die unter Verletzung der Axiome 1-5 aus den

[8] Eine strikte mathematische Äquivalenz würde allerdings gewisse Bedingungen an das Mengensystem Z der „möglichen Welten" fordern, die im Appendix A diskutiert werden. Beispielsweise müsste man eine zu Z zugehörige σ-Algebra Σ definieren, die zusammen mit Z einen sogenannten Massraum bildet. Nur Elementen von Σ könnte dann streng genommen eine Wahrscheinlichkeit zugewiesen werden (und nicht einfach jedem Element von Z). Solche Subtilitäten beginnen eine Rolle zu spielen, falls die Menge Z unendlich viele Elemente enthält, was für unsere Menge Z mit Sicherheit der Fall ist. Beispielsweise ist es nach den wahrscheinlichkeitstheoretischen Axiomen von Kolmogorow unmöglich, eine Lotterie zu veranstalten, in welcher eine beliebige reelle Zahl aus dem Zahlenintervall $[0, 1]$ gezogen wird, wobei jeder Zahl die gleiche Wahrscheinlichkeit zukommt. Mathematisch gesprochen bedeutet das, es gibt keine Wahrscheinlichkeitsverteilung auf der σ-Algebra aller Teilmengen des Intervalls $[0, 1]$, so dass $p([a, b]) = b - a$ für $[a, b] \subseteq [0, 1]$. Man kann also letztlich nur endlichen Teilintervallen eine gleichverteilte Wahrscheinlichkeit zuordnen.

Plausibilitäten anderer Aussagen abgeleitet wurde. Eine solche Handlung müsste man im umgangsprachlichen Sinne als „unlogisch“ oder „irrational“ bezeichnen.

Bayes Theorem

Eines der bedeutendsten Konsequenzen der Bayesianischen Axiome ist ein Theorem, das sich daraus ableiten lässt. Mit dem 5. Axiom gilt für beliebige Aussagen q, r und s

$$p(q \wedge r|s) = p(q|r \wedge s)p(r|s) \ , \tag{2}$$

aber auch

$$p(q \wedge r|s) = p(r|q \wedge s)p(q|s) \ . \tag{3}$$

Daraus folgt sofort Bayes Theorem

$$p(q|r \wedge s) = \frac{p(r|q \wedge s)p(q|s)}{p(r|s)} \ , \tag{4}$$

welches von Reverend Thomas Bayes (1702-1761) gefunden wurde. Die Terme $p(q|s)$ und $p(r|s)$ werden **Prior-Wahrscheinlichkeiten** und der Term $p(q|r \wedge s)$ **Posterior-Wahrscheinlichkeit** genannt.

Die Bedeutung dieses Theorems ist wie folgt: Sei h eine Hypothese, e neues Beweismaterial und k ein Satz von Hintergrundwissen, welches allgemein anerkannt ist. Nun möchte man wissen, wie wahrscheinlich die Hypothese h unter Berücksichtigung des Hintergrundwissens k und des neuen Beweismaterials e. Das heisst man möchte $p(h|e \wedge k)$ kennen. Allerdings ist es oft schwierig, von e direkt auf p zu schliessen. Es ist dagegen häufig einfacher, zu beurteilen, wie wahrscheinlich das Beweismaterial e ist, falls die Hypothese h zutrifft. Und genau diese Information kann mithilfe von Bayes Theorem benutzt werden, um $p(h|e \wedge k)$ zu berechnen. In unserem Fall lautet es

$$p(h|e \wedge k) = \frac{p(e|h \wedge k)p(h|k)}{p(e|k)} \ . \tag{5}$$

Während auf der linken Seite der Gleichung unsere gesuchte Wahrscheinlichkeit ist, so tritt auf der rechten Seite der Gleichung derselbe Ausdruck nochmals auf aber mit h und e vertauscht, d.h. $p(e|h \wedge k)$. Dann gilt es nur noch die Prior-Wahrscheinlichkeiten zu bestimmen. Diese sagen aus, wie wahrscheinlich die Evidenz e bzw. die Hypothese h sind, wenn nur das Hintergrundwissen k verwendet wird. **Damit sagt Bayes Theorem letztlich aus, wie die Wahrscheinlichkeit einer Hypothese h im Lichte neuer Evidenz e verändert wird**, d.h. den Übergang von $p(h|k)$ zu $p(h|e \wedge k)$. Falls nun gilt $p(h|e \wedge k) > p(h|k)$, so sagt man, die Evidenz e bestätigt die Hypothese h. Dies ist genau dann der Fall, falls $p(e|h \wedge k) > p(e|k)$, d.h. falls die Evidenz e unter der Hypothese h wahrscheinlicher erwartet werden kann als ohne die Hypothese h.

Es wird sich als hilfreich erweisen, Bayes Theorem ein wenig umzuformen, da der Term $p(e|k)$ im Nenner gewöhnlich schwer zu berechnen ist. Bezeichnet $\neg h$ der Negation von h, so erhalten wir mit dem 4. Bayesianischen Axiom folgende Relation

$$1 = p(h \vee \neg h|e \wedge k) = p(h|e \wedge k) + p(\neg h|e \wedge k) = \frac{p(e|h \wedge k)p(h|k)}{p(e|k)} + \frac{p(e|\neg h \wedge k)p(\neg h|k)}{p(e|k)} \ , \tag{6}$$

da $h \vee \neg h$ immer wahr ist und somit der gesamte Ausdruck gleich 1 sein muss (Axiom 3). Für die letzte Gleichheit haben wir zweimal Bayes Theorem Gleichung (5) angewendet. Durch Multiplikation mit $p(e|k)$ finden wir schliesslich den allgemeinen Ausdruck

$$p(e|k) = p(e|h \wedge k)p(h|k) + p(e|\neg h \wedge k)p(\neg h|k)\ . \tag{7}$$

Setzen wir diesen Ausdruck in Gleichung (5) ein, erhalten wir den allgemeinen Ausdruck

$$p(h|e \wedge k) = \frac{p(e|h \wedge k)p(h|k)}{p(e|h \wedge k)p(h|k) + p(e|\neg h \wedge k)p(\neg h|k)}\ , \tag{8}$$

den wir im Folgenden für Bayes Theorem verwenden werden.

2.4. Prior-Wahrscheinlichkeiten

Eines der Hauptprobleme im BS ist die Bestimmung der Prior-Wahrscheinlichkeiten $p(h|k)$, wobei h und k grundsätzlich beliebige Aussagen sind. Dieser Ausdruck besagt, wie wahrscheinlich die Hypothese h ist, falls die Aussage k zutrifft. Damit übernimmt k die Rolle des „Hintergrundwissens“ oder der Grundvoraussetzung, von dem wir ausgehen. In philosophischen Kontexten, in denen es um sehr fundamentale Fragen geht, wissen wir unter Umständen fast nichts Sicheres, von dem wir ausgehen könnten. In solchen Fällen kann die Aussage k schrumpfen, bis sie keine Informationen mehr enthält ausser sogenanntes „tautologisches Wissen“, d.h. Wissen, das per Definition wahr ist, wie zum Beispiel „Ein Dreieck hat drei Ecken“. Solches „Hintergrundwissen“ macht keinerlei einschränkende Aussagen, unter welcher die Hypothese h wahr ist.[9] Für den Fall, dass k nur tautologisches Wissen enthält, führen wir die **intrinsische Wahrscheinlichkeit** $p(h)$ (Swinburne 2001, p. 105) durch

$$p(h) = p(h|k) \tag{9}$$

ein. $p(h)$ ist also die Wahrscheinlichkeit, dass die Hypothese h wahr ist unter der Bedingung, dass keine Informationen vorliegen ausser solcher, die bei Definition wahr sind (tautologisches Wissen) und daher unbrauchbar für die Beurteilung, ob die Hypothese h wahr ist oder nicht.

Wie sollen nun die intrinsischen Wahrscheinlichkeiten $p(h)$ und $p(e)$ abgeschätzt werden? Falls es nicht möglich ist, objektive Kriterien anzugeben und zu begründen, warum diese Kriterien gültig sind, wird den intrinsischen Wahrscheinlichkeiten (und schliesslich den Prior-Wahrscheinlichkeiten schlechthin) zwangsläufig subjektives Ermessen anhaften. Zudem wird es kaum je möglich sein, intrinsischen Wahrscheinlichkeiten numerische Werte zuzuweisen. In der Regel wird es nur möglich sein, die intrinsischen Wahrscheinlichkeiten von zwei Hypothesen gegeneinander abzuwägen. Das Problem der intrinsischen Wahrscheinlichkeiten wird weiter in Appendix C behandelt.

Bedeutet das, dass das BS unzuverlässig oder gar unbrauchbar ist? Die Antwort auf diese Frage, falls man sie geben kann, ist tiefgründig und ernüchternd zugleich. Das Problem der Prior-Wahrscheinlichkeiten stellt tatsächlich eine starke Begrenzung der Objektivität des BS dar, besonders wenn man metaphysische Fragestellungen behandelt, über die wir kaum gesichertes Wissen oder Erfahrung haben. Aber haben wir denn überhaupt eine Wahl? Wenn nicht

[9] Die zu k zugehörige Menge von Welten im Sinne von Appendix A ist damit automatisch die Menge Z aller möglichen Welten, da die Aussage k in allen möglichen Welten wahr ist.

das BS, welches rationale Schliessverfahren soll angewendet werden? Man beachte, dass auch die anderen klassichen Schliessverfahren wie Induktion oder Abduktion alle diese subjektiven Elemente gleichermassen enthalten, nur legen sie sie nicht so offen auf den Tisch. Der Mensch wird hier automatisch mit den Grenzen seiner Erkenntnismöglichkeit konfrontiert.

2.5. Rechenbeispiel

Um die in diesem Kapitel eingeführten Konzepte anschaulich zu illustrieren, wird im Folgenden ein einfaches Rechenbeispiel anhand eines medizinischen Tests detailliert durchgeführt.

Ein Mann geht zum Arzt, um sich untersuchen zu lassen. Der Arzt führt nach eingehender Untersuchung einen medizinischen Test durch, wobei das Testergebnis positiv ausfällt. Nun weiss man aus klinischen Testreihen, dass der vom Arzt verwendete Test mit 98 prozentiger Wahrscheinlichkeit positiv ausfällt, falls die vermutete Krankheit vorliegt. Man weiss aber auch, dass der Test bei 4% der Fälle positiv ausfällt, falls die Krankheit nicht vorhanden ist („false positive"). Weiter ist bekannt, dass die Krankheit, um die es sich handeln könnte, ziemlich selten ist, so dass die Wahrscheinlichkeit, dass man die Krankheit in sich trägt, nur 0,4% beträgt. Was kann nun der Arzt über den Gesundheitszustand dieses Mannes aussagen? Aufgrund des durchgeführten Tests kann der Arzt keine *sichere* Aussage über die Gesundheit des Mannes machen, da der Test weder die Anwesenheit noch die Abwesenheit der Krankheit mit absoluter Sicherheit angeben kann, aber mithilfe des BS kann der Arzt die Wahrscheinlichkeit (Plausibilität), dass der Mann wirklich krank ist, zuverlässig abschätzen.

Wir formulieren dazu die Hypothese

$$h = \text{“Der Mann hat die vermutete Krankheit.“} \,,$$

an welcher wir interessiert sind. Die Evidenz e liegt in Form des Tests vor und besteht aus drei Teilaussagen $e = e_1 \wedge e_2 \wedge e_3$, wobei

$$\begin{aligned}
e_1 &= \text{„Der Test ist positiv ausgefallen.“} \,, \\
e_2 &= \text{„Bei Anwesenheit der vermuteten Krankheit fällt der Test generell zu 98\% positiv aus.“} \,, \\
e_3 &= \text{„Bei Abwesenheit der vermuteten Krankheit fällt der Test generell zu 4\% positiv aus.“} \,.
\end{aligned}$$

Das Hintergrundwissen liegt mit

$$k = \text{“Die Krankheit tritt in der Bevölkerung zu 0,4\% auf.“}$$

vor. Dabei sind wir an der Wahrscheinlichkeit interessiert, dass der Mann krank ist unter Berücksichtigung der Evidenz e und des Hintergrundwissens k, d.h. wir möchten den Ausdruck $p(h|e \wedge k)$ bestimmen. Dies geschieht sehr leicht, wenn wir Bayes Theorem in Form von Gleichung (8) benutzen. Dazu bestimmen wir zuerst jede der in dieser Formel auftretenden Wahrscheinlichkeiten. Der Ausdruck $p(e|h \wedge k)$ ist leicht zu berechnen und ergibt 0,98, da h aussagt, dass der Mann krank ist, und der Test unter diesen Bedingungen zu 98% positiv ist. Weiter ist $p(h|k) = 0,004$, da der Mann als Teil der Bevölkerung grundsätzlich eine Chance von 0,4% hat, krank zu sein. Die übrigen beiden Ausdrücke enthalten die Negation $\neg h =$ „Der Mann hat die vermutete Krankheit nicht". Sie ergeben $p(e|\neg h \wedge k) = 0,04$, da der Test im Falle der Abwesenheit die entsprechende Wahrscheinlichkeit hat, positiv zu sein, und $p(\neg h|k) = 1 - p(h|k) = 0,996$

nach Axiom 4. Nehmen wir all diese Wahrscheinlichkeiten zusammen, erhalten wir mittels Bayes Theorem die gesuchte Wahrscheinlichkeit zu

$$p(h|e \wedge k) = \frac{p(e|h \wedge k)p(h|k)}{p(e|h \wedge k)p(h|k) + p(e|\neg h \wedge k)p(\neg h|k)} \tag{10}$$

$$= \frac{0,98 \cdot 0,004}{0,98 \cdot 0,004 + 0,04 \cdot 0,996} \tag{11}$$

$$\simeq 0,09\,. \tag{12}$$

Die Wahrscheinlichkeit, dass der Mann diese Krankheit hat, ist also nur ungefähr 9%, obwohl der Test positiv ausgefallen ist. Der Grund dafür ist die enorme Seltenheit der Krankheit, die als Prior-Wahrscheinlichkeit in unsere Berechnungen einfliesst. Aufgrund des Tests kann der Arzt also nicht auf die Anwesenheit der Krankheit schliessen, sondern müsste dazu weitere Tests durchführen, die mehr Aufschluss über die mögliche Anwesenheit dieser speziellen Krankheit geben.

3. Anwendung des Bayesianischen Schliessverfahrens auf Intelligent Design

Betrachten wir als Beispiel die Entstehung des Lebens. Bis heute ist die Frage völlig ungeklärt, wie das erste Lebewesen entstanden ist. Es ist insbesondere kein natürlicher Mechanismus bekannt, der leblose Materie in Leben überführen kann. Bedeutet das also, dass ein Designer das erste Lebewesen geschaffen hat? Um diese Frage zu beantworten, führen wir im Abschnitt 3.1 das BS zu dieser Fragestellung aus Sicht eines ID Anhängers und aus Sicht eines Naturalisten durch. Ich werde die Argumentationsweisen bewusst relativ vage und allgemein halten, da es mir nicht um eine konkrete, abschliessende Argumentation geht, sondern darum aufzuzeigen, wo im BS möglicherweise unterschiedliche Beurteilungen einfliessen können, die schliesslich zu verschiedenen Resultaten führen. Im Abschnitt 3.2 werde ich die beiden Argumentationsweisen miteinander vergleichen und kommentieren.

3.1. Mögliche Argumentationsweisen

Die zentrale Hypothese sei $h =$ „Das erste Lebewesen wurde von einem Designer erschaffen.“ Diese Aussage soll mithilfe der folgenden Aussage $e =$ „Die erste denkbare Zelle muss ungeheuer komplex gewesen sein“ untermauert werden. Weiter steht uns das Hintergrundwissen k zur Verfügung, welches unser heutiges naturwissenschaftliches Wissen umfasst. Wir wollen nun eine Abschätzung für $p(h|e \wedge k)$ vornehmen. Das ist die Wahrscheinlichkeit, dass das erste Lebewesen von einem Designer erschaffen wurde, wobei wir annehmen, dass es ungeheuer komplex war, und für die Beurteilung unser heutiges naturwissenschaftliches Wissen voraussetzen. Um die Abschätzung durchzuführen, gehen wir von Bayes Theorem in der Form von Gleichung (8) aus:

$$p(h|e \wedge k) = \frac{p(e|h \wedge k)p(h|k)}{p(e|h \wedge k)p(h|k) + p(e|\neg h \wedge k)p(\neg h|k)}\,. \tag{13}$$

Diese Formel wird für unsere Diskussion sogar noch etwas klarer, wenn wir sie noch ein wenig umformen zu

$$p(h|e \wedge k) = \frac{1}{1 + \frac{p(e|\neg h \wedge k)p(\neg h|k)}{p(e|h \wedge k)p(h|k)}} \,. \tag{14}$$

Damit wird deutlich, dass die gesuchte Wahrscheinlichkeit $p(h|e \wedge k)$ nur vom Verhältnis

$$\frac{p(e|\neg h \wedge k)p(\neg h|k)}{p(e|h \wedge k)p(h|k)} \tag{15}$$

abhängt. Wir müssen also nicht die einzelnen Ausdrücke in Gleichung (14) bestimmen, sondern es reicht, ihren relativen Wert zueinander zu abzuschätzen. Je kleiner das Verhältnis (15) wird, desto höher die Wahrscheinlichkeit $p(h|e \wedge k)$. Eine solche Abschätzung soll nun beispielhaft zuerst aus der Sichtweise eines ID Anhängers und dann aus der Sichtweise eines Naturalisten durchgeführt werden. Die folgende Argumentationsweisen sind rein hypothetischer Natur und wurden bewusst so gewählt, dass sie gewisse Punkte sehr deutlich machen.

ID Anhänger

Der ID Anhänger hat i.d.R. keine Berührungsängste mit dem metaphysischen Konzept des Designers. Meistens glauben diese Menschen an Gott und identifizieren ihn mit dem Designer. Ich setze diese Grundposition im Folgenden voraus. So eine Person wird das Verhältnis

$$\frac{p(\neg h|k)}{p(h|k)} \tag{16}$$

vermutlich irgendwo in der Mitte zwischen 0 und 1 anordnen. Für diese Person ist das Hintergrundwissen k gut mit der Position vereinbar, dass ein Designer (= Gott) existiert (hat), der das erste Lebewesen erschaffen hat. Zwar mag k nicht direkt dafür sprechen, aber es spricht auch nicht dagegen.

Bei Betrachtung des verbleibenden Terms

$$\frac{p(e|\neg h \wedge k)}{p(e|h \wedge k)} \tag{17}$$

fällt die Komplexität e des Lebens ins Gewicht. Während sich der ID Anhänger sehr gut vorstellen kann, dass der Designer komplexes Leben erschaffen hat, d.h. $p(e|h \wedge k)$ ist nahe bei 1 – schliesslich ist ein Designer so definiert, dass er Schöpferisch tätig ist und ein sehr grosser Designer kann auch sehr komplexe Dinge erschaffen – so wird der ID Anhänger seine Argumente darlegen, dass $p(e|\neg h \wedge k)$, die Wahrscheinlichkeit, dass es komplexes Leben gibt, das nicht von einem Designer erschaffen wurde, sehr klein, ja gerade verschwindend klein sein muss. Alles, was er weiss, d.h. k, spricht dagegen, dass so etwas komplexes wie Leben von alleine entstehen könnte. Er setzt daher $p(e|\neg h \wedge k) \ll 1$ oder mindestens $p(e|\neg h \wedge k) \ll p(e|h \wedge k)$.

Insgesamt ergibt sich also für den ID Anhänger die Bilanz

$$p(h|e \wedge k) = \frac{1}{1 + \frac{p(e|\neg h \wedge k)p(\neg h|k)}{p(e|h \wedge k)p(h|k)}} \sim 1 \,. \tag{18}$$

Das heisst, es ist angesichts der Fakten sehr wahrscheinlich, dass die Hypothese h wahr ist, d.h. dass ein Designer Leben erschaffen hat. Der treibende Beitrag für dieses Ergebnis ist der winzig kleine Wert der naturalistischen Lebensentstehung $p(e|\neg h \wedge k)$.

Naturalist

Der Naturalist wird freilich zu einem anderen Ergebnis kommen. Das hat zwei Hauptgründe. Zum einen wird der Naturalist der Analyse des ID Anhängers kaum zustimmen. Zwar ist es gut möglich, dass auch er einräumt, dass $p(e|\neg h \wedge k) \ll p(e|h \wedge k)$, d.h. eine naturalistische Entstehung des Lebens sehr unwahrscheinlich ist. Auf der anderen Seite wird er dem ID Anhänger aber kaum zustimmen, dass der Term

$$\frac{p(\neg h|k)}{p(h|k)} \tag{19}$$

irgendwo in der Mitte zwischen 0 und 1 ist. Für einen Naturalisten könnte es geradezu grotesk anmuten, anzunehmen, die Existenz eines Gottes sei apriori etwas „Natürliches“ oder „Wahrscheinliches“. Der Naturalist ist aufgrund seines Hintergrundwissens k überzeugt, dass die Welt ohne die Hilfe eines Gottes funktioniert. Bisher habe die Forschung den reduktionistischen Ansatz immer bestätigt. Es sei zwar naheliegend, den Faktor $p(e|h \wedge k)$ irgendwo in der Mitte zwischen 0 und 1 anzusetzen, denn wir wissen heute aus der Gentechnik, dass intelligente Wesen in der Lage sind, biologisch kreativ tätig zu sein. Auf der anderen Seite wissen wir auch, dass es bei der Entstehung des Lebens noch keine Menschen gegeben hätte und von anderer Intelligenz fehlt bislang jede Spur. Von daher wird der Naturalist möglicherweise zum Schluss kommen, dass

$$\frac{p(\neg h|k)}{p(h|k)} \gg 1\,. \tag{20}$$

Ja die Vorstellung eines Gottes könnte ihm so abwegig erscheinen, dass er sie sogar gegen die Komplexität des Lebens auf die Waage wirft. Er wird also

$$\left(\frac{p(\neg h|k)}{p(h|k)}\right)^{-1} \ll \frac{p(e|\neg h \wedge k)}{p(e|h \wedge k)} \ll 1 \tag{21}$$

setzen, was zum Ausdruck bringt, dass aus seiner Sicht die natürliche Entstehung des Lebens zwar sehr unwahrscheinlich ist, aber die Existenz eines Gottes noch unvergleichlich unwahrscheinlicher. Setzen wir beide Verhältnisse aus Gleichung (21) in Gleichung (14) ein, so führt das schliesslich zu

$$p(h|e \wedge k) = \frac{1}{1 + \frac{p(e|\neg h \wedge k)p(\neg h|k)}{p(e|h \wedge k)p(h|k)}} \ll 1\,. \tag{22}$$

Für den Naturalisten erscheint also selbst angesichts der Komplexität des Lebens die Hypothese h sehr unwahrscheinlich. Die treibende Kraft hinter diesem Schluss ist die winzige Prior-Wahrscheinlichkeit $p(h|k)$, die sich für ihn aus seinem Weltverständnis ergibt und die sogar das sehr unwahrscheinliche Ereignis einer naturalistischen Lebensentstehung überwiegt.

Aber nicht alle Naturalisten würden wohl so argumentieren wie im vorhergehenden Abschnitt. Es gibt einen zweiten Hauptgrund, warum sie ID ablehnen. Ein Naturalist könnte nämlich durchaus einräumen, dass der ID Anhänger die Wahrscheinlichkeit $p(h|e \wedge k)$ richtig abgeschätzt hätte, dass aber dieser Ausdruck für die Frage nach der Entstehung des Lebens irrelevant sei. Einerseits könnte es gut sein, dass sich in der Zukunft herausstellen wird, dass die Grundvoraussetzung e, wonach das erste Leben sehr komplex war, falsch ist. Des Weiteren könnte er beanstanden, dass unser heutiges Wissen k viel zu unvollständig ist, als dass eine zuverlässige Aussage über die Entstehung des Lebens gemacht werden könne. Damit wäre er vorerst aus dem Schneider.

Für diese letzte Argumentationsweise gibt es noch einen weitere Variante. Ein Naturalist könnte nicht nur zur Einsicht kommen, dass der ID Anhänger die Wahrscheinlichkeit $p(h|e \wedge k)$ richtig abgeschätzt hätte, sondern auch zustimmen, dass diese Wahrscheinlichkeit relevant sei, da er nicht glaube, dass in Zukunft ein naturalistischer Mechanismus gefunden würde, der den Übergang von lebloser zu lebender Materie erklären könne. Wie kann ein solcher Naturalist seine Position aufrecht erhalten? Er könnte beispielsweise anführen, dass nicht spezifiziert worden sei, was unter 'Designer' in der Hypothese h genau gemeint sei. Dieser Designer muss ja nicht zwingenderweise super-naturalistischer Art sein, sondern könnte beispielsweise (naturalistische) Aliens darstellen, unseren Planeten mit Leben infiziert hätten.[10] Oder er fängt an, von einem Multiversum zu erzählen, dass so viele Planeten beinhalten würde, dass selbst sehr unwahrscheinliche Ereignisse irgendwann mal passieren (müssen). Mit diesem letzten Punkt hätte der Naturalist das Argument in eine ganz andere Richtung gelenkt, da im naturwissenschaftlichen Hintergrundwissen k ein Multiversum keineswegs als gesicherte Tatsache enthalten ist. Der Naturalist würde aber seinerseits deutlich machen, dass er zur Beantwortung der Frage nach der Entstehung des Lebens nicht nur das Beweismaterial e, sondern möglicherweise weitere Diskussionsgrundlagen hinzunehmen möchte, wie z.B. e_2 = „Es gibt ein Multiversum." Unter diesen Umständen würde gelten $p(e|\neg h \wedge e_2) \sim 1$, aber man müsste sich damit auseinandersetzen, dass die Prior-Wahrscheinlichkeit für ein Multiversum $p(e_2|k)$ möglicherweise sehr klein ist.

3.2. Diskussion

Ein Vergleich der beiden hypothetischen Argumentationsmuster des ID Anhängers und des Naturalisten lassen einige interessante Beobachtungen zu:

- Da ein Ausdruck $p(h|e \wedge k)$ nur im Rahmen von e und k gültig ist, ist es wichtig, die Aussagen h, e und k sinnvoll zu wählen. Falls e und k ungeschickt gewählt werden, kann es zwar unter Umständen möglich sein, dem Ausdruck $p(h|e \wedge k)$ einen Wert nahe bei 1 zuzuweisen, dies ist jedoch wenig zielführend, wenn der gesamte Ausdruck für sich genommen wenig aussagekräftig ist. Dem Ausdruck $p(h|e \wedge k)$ keine Relevanz zuzusprechen, stellte ja eine Strategie des Naturalisten dar. Möchte man eine Thematik abschliessend behandeln, ist es zudem notwendig, *sämtliche* Aussagen (Fakten, Evidenz) e zu berücksichtigen, die für die Fragestellung relevant sein könnten. Für die Entstehung des Lebens müsste man also z.B. auch die Möglichkeit anderer Planeten und gar eines Multiversums in Betracht ziehen. Vergisst man eine solche Evidenz e, läuft man Gefahr, dass der Ausdruck $p(h|e \wedge k)$ als nicht relevant eingestuft wird.

- Bei der Anwendung des BS ist selbst bei gut gewählten h, e, und k nicht garantiert, dass alle zur selben Schlussfolgerung kommen. Die Argumentationsweisen des ID Anhängers und des Naturalisten demonstrieren das Ausmass, in welchem subjektives Ermessen das BS beeinflussen kann. Wie in Abschnitt 2.4 angesprochen, ist das nicht in erster Linie ein Mangel des BS, sondern eine grundsätzliche Schwierigkeit der behandelten Fragestellung.

[10] Eine unmittelbare Reaktion eines ID Vertreters auf diese Argumentationsweise könnte die Frage sein, wie denn die Aliens entstanden seien. Der Naturalist könnte sich z.B. mit der vorläufigen Antwort behelfen, dass die Aliens eine Lebensform darstellen würden, deren Entstehung anders als irdisches Leben unter geeigneten Bedingungen möglicherweise viel wahrscheinlicher sei.

Bei der Frage nach Gottes Existenz gehen nun mal die Meinungen auseinander und es ist möglich, dass solche Fragen auf objektive Weise für uns Menschen nicht entscheidbar sind. Das BS zeigt aber so klar wie nur möglich, inwiefern solche Meinungsverschiedenheiten die Schlussfolgerung beeinflussen. Der abduktive Schluss (Schluss auf die beste Erklärung) beispielsweise wäre bei Anwendung auf diese Fragestellung nicht weniger subjektiv, nur dass bei diesem Schluss der ID Anhänger und der Naturalist einfach zu unterschiedlichen Ergebnissen kommen würden, ohne dass ersichtlich wäre, wo diese Unterschiede entstehen. Was für den einen der Schluss auf die beste Erklärung ist, muss nicht auch für den anderen so sein.

- Die Argumentationsmuster machen deutlich, dass die Prior-Wahrscheinlichkeiten einen sehr grossen Einfluss auf die Meinungsbildung haben können. Ob man an Gott glaubt oder nicht, kann bei dieser Fragestellung für das Ergebnis entscheidend sein. Die Weltanschauung und „Ontologie" des Betrachters fliessen in die Prior-Wahrscheinlichkeiten. Der naturwissenschaftliche Aspekt der Fragestellung ist vollständig im Ausdruck $p(e|\neg h \wedge k)$ enthalten und bildet damit nur einen von wenigstens vier Beiträgen.

- In den Beispielen wurden die Prior-Wahrscheinlichkeiten nach Gutdünken festgelegt. Sind sie bei solchen Fragestellungen nur reine Meinungssache? Swinburne ist der Ansicht, dass Prior-Wahrscheinlichkeiten auch nach rationalen Gesichtspunkten abgeschätzt werden können. Seine Position ist im Appendix C skizziert. Ein Beispiel dafür hat Swinburne mit seinem Buch „The Existence of God" gegeben, in welchem er versucht, mithilfe des BS darzulegen, dass die Hypothese $h =$ „Gott existiert." eine Wahrscheinlichkeit $p(h|e) > 0,5$ hat. Als Beweismaterial e nimmt Swinburne hauptsächlich Fakten, die in unserem Beispiel im Hintergrundwissen k enthalten wären. Ich bin zwar nicht der Ansicht, dass jeder seiner Analyse zustimmen müsste, aber sein Buch ist für mich ein Hinweis, dass es sich durchaus lohnen kann, über Prior-Wahrscheinlichkeiten nachzudenken.[11] Je mehr man über sie sagen kann, desto zuverlässiger wird die Schlussfolgerung aus dem BS.

- Wie bereits erwähnt, ist der naturwissenschaftliche Aspekt unserer ID Fragestellung allein im Term $p(e|\neg h \wedge k)$ enthalten. Je genauer unser Informationsgehalt e über die Entstehung des Lebens, desto präziser kann dieser Term abgeschätzt werden. Sollte e beispielsweise etwas über die Randbedingungen auf der frühen Erde und den Grad der Komplexität der ersten Zelle aussagen, so könnte eine Abschätzung von $p(e|\neg h \wedge k)$ möglicherweise mittels statistischer Methoden erfolgen.

4. Zusammenfassung

Im ersten Kapitel wurde das BS konzeptionell eingeführt. Es wurde dargelegt, welche Bedeutung die mit dem BS verbundenen induktiven Wahrscheinlichkeiten haben und wie damit gerechnet werden kann. Das BS ist im Prinzip eine Rechenvorschrift, wie die Plausibilität von Aussagen (z.B. einer Theorie oder Hypothese) von den Plausibilitäten anderer Aussagen abgeleitet werden können. Es beruht auf fünf Axiomen, die so gewählt wurden, dass das BS kohärent

[11] Eine Rezension dieses Buches von mir findet man unter http://www.wort-und-wissen.de/index2.php?artikel=info/rezens/b21.html. Ich würde in der Tat nicht jedem Aspekt seiner Argumentation folgen.

ist. Das bedeutet, dass es zu Ergebnissen führt, die mit der mathematischen Wahrscheinlichkeitstheorie (und damit mit der Logik) übereinstimmen. Wer also seine Schlussfolgerung auf diese Axiome gründet, argumentiert rational, und wer eines der Axiome missachtet, argumentiert möglicherweise nicht rational und kann zu Fehlschlüssen kommen. Aus den Axiomen lässt sich Bayes Theorem ableiten, welches ermöglicht, die Plausibilität einer Hypothese h unter Berücksichtigung bestimmter Fakten e zu bestimmen, indem untersucht wird, wie plausibel das Auftreten der Fakten e unter Annahme der Hypothese h wäre. Allerdings treten dabei Prior-Wahrscheinlichkeiten auf, die in vielen Fällen nicht leicht zu bestimmen sind und die das BS letztlich ein Stück weit der Subjektivität des Betrachters unterwerfen. Das ist jedoch nicht in erster Linie eine Schwäche des BS, sondern eine Konsequenz unseres begrenzten Wissens. Das BS kann nicht Informationen erzeugen, die vorher noch nicht da waren, aber das BS zeigt, wie man mit vorhandenen Informationen richtig argumentiert.

Neben seiner Kohärenz, liegt die Stärke des BS in seiner klaren Struktur. Man wird durch das Einhalten der 5 Axiome quasi dazu gezwungen, korrekt zu argumentieren, und durch das Betrachten der einzelnen Beiträge in Bayes Theorem kann ersichtlich werden, warum zwei Menschen bei gleicher Ausgangslage nicht zum selben Resultat kommen. Anhand der Prior-Wahrscheinlichkeiten sieht man zudem, wie die Weltanschauung (bzw. Ontologie) des Betrachters die Schlussfolgerung beeinflussen kann. Je besser man die Prior-Wahrscheinlichkeiten im Griff hat, desto zuverlässiger wird das BS. Trotz der Stärken des BS ist es wertlos, wenn es nicht richtig angewendet wird. Das zeigten die hypothetischen Argumentationsweisen des ID Vertreters und des Naturalisten. Wer es aber versteht, das BS korrekt und geschickt anzuwenden, kann es mit grossem Gewinn und viel Klarheit auf die Ursprungsfrage anwenden.

Literatur

Chalmers, A.F. (1999) *What is this thing called Science?*, Third edition, Mcgraw-Hill Higher Education.

Cowan G. (1998) *Statistical Data Analysis*, Oxford University Press.

Rautenberg W. (2008) *Einführung in die Mathematische Logik*, 3. Auflage, Vieweg+Teubner.

Swinburne R. (2001) *Epistemic Justification*, Oxford University Press.

Swinburne R. (2004) *The Existence of God*, Second edition, Oxford University Press.

Zoglauer T. (2008) *Einführung in die formale Logik für Philosophen*, 4. Auflage, Vandenhoeck & Ruprecht.

Appendices

A. Aussagen als Mengen interpretieren

Im Haupttext dieses Essays wurde das BS mithilfe von „Aussagen“ formuliert. Es gibt jedoch die Möglichkeit, Aussagen als mathematische Mengen zu interpretieren, was einige Vorteile haben kann. Man begibt sich damit auf mathematischen Boden, indem das BS auf die Grundlage der Mengenlehre gestellt wird, und wird damit zur höchstmöglichen Genauigkeit gezwungen. Die beiden Zugänge sind aber im Prinzip völlig äquivalent.

Die Formulierung mittels Mengen beruht darauf, dass jede Aussage einer gewissen Menge zugeordnet wird und umgekehrt. Die Elemente dieser Mengen sind „**mögliche Welten**“.[12] Diese „möglichen Welten“ sind gedachte Möglichkeiten, wie die Wirklichkeit sein könnte. Grundsätzlich sind alle Möglichkeiten (Welten) zugelassen, die nicht in sich widersprüchlich sind (daher die Präzisierung „möglich“). „Welt“ bedeutet hier also nicht Planet oder Universum, sondern ist umfassender. Denkbare Wirklichkeiten gibt es unendlich viele und nur eine einzige ist davon ist effektiv realisiert (also wirklich). Eine Welt grenzt quasi das Existierende vom Nicht-existierenden ab. (Es ist aber möglich, dass eine gewisse denkbare Welt identisch mit dem Universum ist. In diesem Fall wäre das Universum alles, was existiert. Es gäbe keine anderen Universen und auch keine übernatürlichen Wesen.) Eine Aussage entspricht nun der Menge aller möglichen Welten, in welcher die Aussage wahr ist. Die Vereinigung aller möglichen Welten bezeichne ich als Menge Z.[13] Jede Aussage entspricht also einer Teilmenge dieser Menge Z.

Zum Beispiel bezieht sich die Aussage $q =$ „Menschen stammen von affenartigen Vorfahren ab“ auf die Menge Q aller logisch denkbaren Welten (egal wie abstrakt sie sein mögen), in welchen Menschen von affenartigen Vorfahren abstammen.[14] Ein anderes Beispiel ist die zugehörige Menge zu in sich widersprüchlichen Aussagen. Diese Menge muss der leeren Menge $\emptyset$ entsprechen, da es keine denkbar mögliche Welt gibt, in welcher eine in sich widersprüchliche Aussage wahr sein könnte.

Die Interpretation von Aussagen als Mengen hat den Vorteil, dass sie die menschliche Sprache in ein mathematisches Objekt übersetzt, das idealerweise eindeutig definiert und dadurch objektiv zugänglich ist. Die Denkweise in Mengen zwingt zudem dazu, sich die Bedeutung von Begriffen genau zu überlegen. Ein Begriff ist streng genommen sogar erst dann ganz genau definiert, wenn man die Menge der Welten angibt, in welchen der Begriff vorkommt. Denn damit

[12] Solche möglichen Welten kommen in der Philosophie auch in anderen Kontexten vor und sind z.B. ein Grundbestandteil der modalen Logik (siehe z.B. Zoglauer 2008, Kapitel 8).

[13] Es wird hier stillschweigend angenommen, dass diese Menge existiert und beispielsweise der Zermelo-Fraenkel Axiomatik genügt. Sollte sie nicht existieren, so definiere man einfach eine kleinere Menge Z von „Welten“. Wird diese Menge Z allerdings zu klein, läuft man Gefahr, dass man gewisse „Welten“, die denkbar wären, nicht in Betracht zieht und daher mit dem BS eine unsorgfältige Analyse macht. Dies konnte z.B. dann der Fall sein, falls ein Naturalist in der Menge von „Welten“ keine „Welten“ zulässt, die einen Gott beinhalten, oder wenn ein ID-Vertreter nicht alle möglichen Szenarien für eine naturalistische Entstehung des Lebens und der Artenvielfalt in Betracht zieht. Beides kann gleichermassen verhängnisvoll sein. Z sei im Folgenden einfach die Menge der zugelassenen möglichen Welten. Sie entspricht dem, was in der Philosophie manchmal als „Diskursuniversum“ bezeichnet wird.

[14] Aussagen werden im Folgenden immer als kleine Buchstaben dargestellt, während ihre dazugehörigen Mengen mit den entsprechenden grossen Buchstaben bezeichnet werden. In diesem Beispiel also die Aussage q und die dazugehörige Menge Q.

ist für jeden nur denkbaren Fall eindeutig entschieden, ob der Begriff für ein gewisses Objekt zutrifft oder nicht. Wie ist beispielsweise der Begriff „Leben“ genau definiert? Die Bedeutung wird dann eindeutig klar (für alle auch nur denkbaren Fälle), wenn ich mir alle möglichen Welten vorstelle und genau angebe, in welchen von diesen Welten Leben vorkommt und in welchen nicht. Ist eine gigantische kosmische Staubwolke ein Lebewesen? Das hängt davon ab, ob es in einer fiktiven Welt, in welcher es nur diese kosmische Staubwolke gibt und sonst nichts, Leben gibt oder nicht. Allerdings wird es in der Praxis keinesfalls möglich sein, all diese Welten anzugeben, nur schon, weil es unendlich viele gibt und man zur genauen Beschreibung einer gewissen möglichen Welt unendlich viele Angaben machen muss. Daher ist man letztlich doch auf eine sprachliche Beschreibung (und den gesunden Menschenverstand) angewiesen.

Auch die Verknüpfung von Aussagen mittels '$\vee$' und '$\wedge$' wird mathematisch sehr durchsichtig, falls man Aussagen als Mengen interpretiert. Hat man zwei Aussagen q und r, so bedeuten die Verknüpfungen $s_1 = q \vee r$ und $s_2 = q \wedge r$ einfach die Vereinigung ('$\cup$') respektive die Schnittmenge ('$\cap$') der entsprechenden Mengen der Aussagen p und q, d.h. $S_1 = Q \cup R$ und $S_2 = Q \cap R$.

Den Ausdruck $p(q|r)$ kann man auch im Sinne von Mengen interpretieren. Er bedeutet nichts anderes, als dass zur Auswertung der Wahrscheinlichkeit für die Aussage q, die Menge der zugelassenen möglichen Welten Z auf die Menge R bezüglich der Aussage r eingeschränkt wird. Der Anteil der zur Aussage q gehörigen Menge Q, der nicht im Durchschnitt $Q \cap R$ enthalten ist, trägt nichts zur Wahrscheinlichkeit $p(q|r)$ bei. Die Auswertung von $p(q|r)$ wird quasi auf alle möglichen Welten innerhalb von R eingeschränkt. Falls Q teilweise darüber hinausgeht, ist das irrelevant.

B. BS in der physikalischen Datenanalyse

Wie in Abschnitt 2.2 angesprochen, wird das BS in der Physik im Rahmen der Datenanalyse angewendet (siehe z.B. Cowan 1998, Kapitel 6.13). Im Folgenden wird das entsprechende Verfahren kurz dargestellt und kritisch diskutiert.

Sei h ein naturwissenschaftliches Modell, welches von den Parametern $\theta_1, \theta_2 \ldots$ abhängt, d.h. man hat eine ganze Familie von Hypothesen $h(\theta_1, \theta_2 \ldots)$. Nun möchte möchte man anhand einer Messung e bestimmen, welches der Hypothesen $h(\theta_1, \theta_2 \ldots)$ die plausibelste ist. Man möchte also die freien Parameter $\theta_1, \theta_2 \ldots$ so bestimmen, dass das Modell h am besten zu den Daten e passt, d.h. $p(h(\theta_1, \theta_2 \ldots)|e \wedge k)$ soll für diese Wahl von $\theta_1, \theta_2 \ldots$ maximal werden. Das Hintergrundwissen k wird dabei entweder als tautologisches Wissen angenommen oder fliesst im Sinne von „physikalischer Erfahrung“ ein und kann z.B. den Parameterbereich eingrenzen (siehe unten), in welchem man die plausibelste Lösung vermutet. Den gesuchten Ausdruck berechnet man schliesslich mithilfe von Bayes Theorem, indem man die freien Parameter $\theta_1, \theta_2 \ldots$ so bestimmt, dass der Ausdruck

$$p(h(\theta_1, \theta_2 \ldots)|e \wedge k) = \frac{p(e|h(\theta_1, \theta_2 \ldots) \wedge k)p(h(\theta_1, \theta_2 \ldots)| \wedge k)}{p(e| \wedge k)} \tag{23}$$

maximal wird. Da $p(e| \wedge k)$ nicht von den freien Parametern $\theta_1, \theta_2 \ldots$ abhängt, kann man diesen Term getrost gleich 1 setzen. Das hat keinen Einfluss auf die gesuchten Werte der Parameter, bei welchen der Ausdruck 23 maximal wird, d.h. das Maximum bleibt an derselben Stelle. Diese Vorgehensweise ist in der Naturwissenschaft sehr nützlich, da es im Gegensatz zur Diskussion über Intelligent Design im Kapitel 3 hier gewöhnlich möglich ist, den Ausdrücken $p(e|h \wedge k)$ für gegebenes e und gegebenes $h(\theta_1, \theta_2 \dots)$ präzise numerische Werte in Abhängigkeit der freien Parameter $\theta_1, \theta_2 \dots$ zuzuweisen, d.h. man hat für ein gegebenes Modell für alle möglichen freien Parameter konkrete Vorhersagen für den Vergleich mit den Daten.

Das Problem ist nun aber häufig, dass das Hintergrundwissen k kaum etwas über die Parameter $\theta_1, \theta_2, \dots$ aussagt. In diesem Fall ist nicht klar, wie die Prior-Wahrscheinlichkeiten für $p(h(\theta_1, \theta_2, \dots)|k)$ in Abhängigkeit der Parameter gewählt werden müssen. Sollte das Hintergrundwissen gar keine Aussage über die Parameter $\theta_1, \theta_2 \dots$ machen, so ist es naheliegend anzunehmen, dass der Prior-Wahrscheinlichkeit für alle Parameterwerte denselben Wert zukommt, d.h. die Wahrscheinlichkeit für die Parameterwerte sind gleichverteilt. Nun ist es aber so, dass die Parameter für viele Theorien grundsätzlich beliebige Werte annehmen können, d.h. von $-\infty$ bis $+\infty$ ist im Prinzip jeder Wert zugelassen. Wenn aber allen Werten von $-\infty$ bis $+\infty$ eine gleiche Wahrscheinlichkeit zukommt, muss diese Wahrscheinlichkeit Null sein, da sonst bei der Aufsummierung aller möglicher Parameterwerte die Gesamtwahrscheinlichkeit grösser als 1 ist und damit das dritte Axiom verletzen würde. Aus diesem Grund müssen in der Praxis eine obere und untere Schranke für die Parameterwerte gewählt werden. Diese Schranken werden aber mehr oder weniger willkürlich gewählt und enthalten daher letztlich ein subjektives Element. Dieses subjektive Element ist einer der Hauptgründe, warum die Bayesianische Statistik unter Experten z.T. umstritten ist. In der „frequentist statistics“ wäre ein solcher Schritt nicht erforderlich. Es sei aber angemerkt, dass dieses Problem auch von der konkreten Form von $p(e|h(\theta_1, \theta_2 \dots) \wedge k)$ abhängig ist. Falls beispielsweise $p(e|h(\theta_1, \theta_2 \dots) \wedge k)$ für $|\theta_i| \to \infty$, $i = 1, 2, \dots$, schnell genug abfällt, kann die genaue Wahl der Ober- und Unterschranken irrelevant sein, da die Posterior-Wahrscheinlichkeit auch normalisierbar bleibt, falls die Prior-Wahrscheinlichkeit überall konstant angenommen wird.

Eine weitere Entscheidungsfreiheit betrifft die Art und Weise, wie die Hypothese $h(\theta_1, \theta_2 \dots)$ parametrisiert wird. Es macht z.B. einen Unterschied, ob θ_1 als freier Parameter betrachtet wird oder $\tilde{\theta}_1 = 1/\theta_1^2$. Wähle ich nämlich als Prior-Wahrscheinlichkeit den einen der beiden Parameter als gleichverteilt, so ist es der andere nicht. Damit hat die Wahl der Parameter einen Einfluss auf die Prior-Wahrscheinlichkeiten. Dieses Problem besteht im Gegensatz zum Normalisierungsproblem (siehe oben) unabhängig von der konkreten Form von $p(e|h(\theta_1, \theta_2 \dots) \wedge k)$.

C. Intrinsische Wahrscheinlichkeiten

In diesem Appendix soll das Konzept der intrinsischen Wahrscheinlichkeiten näher beleuchtet werden. Gibt es denn einen korrekten, objektiven Weg, um einer beliebigen Aussage eine intrinsische Wahrscheinlichkeit zuzuweisen? Wie bereits in Fussnote 6 angesprochen zeigt sich Swinburne diesbezüglich optimistisch. Wenn es auch nicht möglich sein dürfte, genaue numerische Werte für intrinsische Wahrscheinlichkeiten zu bestimmen, so dürfte dennoch die Grössenordnung oder wenigstens die relative Grösse zu einer anderen intrinsischen Wahrscheinlichkeit objektiv begründet werden. Hinsichtlich der Objektivität der intrinsischen Wahrscheinlichkeit schreibt Swinburne (2001, p. 122):

> If science is really a rational enterprise in the sense that certain evidence really does make one hypothesis more probable and another hypothesis less probable, and

> so there are indeed correct criteria of inductive inference, of what is evidence for what and how strong particular evidence is, there must be principles of probability additional to the Bayesian axioms, and in particular a priori principles for ascribing intrinsic probabilities. [...] But, if there are no such a priori criteria, we should give up studying science, history, and every other subject of university study.

Da Wissenschaft aber offenbar ein sehr erfolgreiches Unternehmen ist und es unter Wissenschaftlern nur wenige Grundsatzdiskussionen gibt, was eine Theorie wahrscheinlicher als eine andere Theorie macht, scheint es offenbar gewisse Prinzipien zu geben, denen sich die meisten Wissenschaftler bewusst oder unbewusst angeschlossen haben. Ein solches Prinzip könnte zum Beispiel „Occam's razor“ sein, das in der Wissenschaft breite Anwendung findet und zum Selektionieren von Theorien geeignet ist. Swinburne hat nun den Versuch unternommen, konkrete Prinzipien zu formulieren, welche die intrinsische Wahrscheinlichkeit einer Hypothese (bzw. Aussage) bestimmen. Er unterscheidet dazu zwei Bereiche: Den Geltungsbereich („scope“) einer Aussage und die Einfachheit („simplicity“) einer Aussage. Grundsätzlich gilt, je kleiner der Geltungsbereich und je grösser die Einfachheit einer Aussage, desto höher ist ihre intrinsische Wahrscheinlichkeit. (Swinburne 2001, p. 82) Dies soll im Folgenden erläutert werden:

Der Geltungsbereich einer Hypothese ist dadurch definiert, wie viel sie über die Welt aussagt. Je mehr Behauptungen man macht und je detaillierter diese Behauptungen sind, desto grösser ist die Wahrscheinlichkeit, dass die Behauptungen Fehler enthalten. Es gibt keinen präzisen Weg, den Geltungsbereich einer Aussage zu „messen“, aber der Geltungsbereich zweier Aussagen kann miteinander verglichen werden. Grundsätzlich gilt, dass falls eine Aussage eine andere enthält, aber nicht umgekehrt, so hat die erstere Aussage einen grösseren Geltungsbereich als die zweite.

Die Einfachheit einer Theorie bekommt nach Swinburne grösseres Gewicht als der Geltungsbereich. Aus diesem Grund widmet er diesem Thema in seinem Buch etwa 20 Seiten (Swinburne 2001, p. 83 ff.), während er dem Geltungsbereich vielleicht eine halbe Seite widmet (p. 82). Das liegt auch daran, dass es gar nicht so einfach ist, in Worte zu fassen, was eine Theorie einfach macht und wie man diese Einfachheit messen soll. Swinburne hält fest (p. 87): „One hypothesis is simpler than another if (and only if) the simplest formulation of the former is simpler than the simplest formulation of the latter.“ Dann nennt Swinburne sechs Facetten von Einfachheit:

- Einfachheit bezüglich der Anzahl an „Dingen“ die postuliert werden. Falls zwei Hypothesen einen Befund gleichermassen erklären, dann soll die Hypothese bevorzugt werden, die weniger „Dinge“ postuliert.

- Einfachheit bezüglich der „Arten“ von Dingen. Falls zwei Hypothesen einen Befund gleichermassen erklären und gleich viele „Dinge“ postulieren, dann soll die Hypothese bevorzugt werden, die am wenigsten verschiedene Dinge postuliert.

- Einfachheit im Sinne von leichter Verständlichkeit. Sollte eine Hypothese einen Begriff enthalten, der wiederum zuerst erklärt werden muss, damit er verständlich ist, so ist sie weniger einfach als eine gleich einfache Formulierung der Hypothese, die auf solche Begriffe verzichten kann.

- Einfachheit im Sinne der Anzahl „Naturgesetze“, die sie beinhaltet. Eine Hypothese, welche ein paar wenige Naturgesetze enthält ist einfacher als eine Hypothese, die viele Naturgesetze enthält.

- Einfachheit bezüglich der Anzahl der Variablen, die durch ein Naturgesetz in Beziehung zueinander gesetzt werden. Eine Hypothese, welche Naturgesetze enthält, die einige wenige Variablen zueinander in Beziehung setzen, ist einfacher als eine Hypothese mit Naturgesetzen, die viele Variablen miteinander in Beziehung setzen.

- Einfachheit im Sinne mathematischer Formulierung. Eine mathematisch einfachere Formulierung ist einfacher. Dabei ist eine mathematische Formulierung umso einfacher, je weniger Terme sie enthält und je einfachere Grössen die Terme beinhalten, wobei eine Grösse X einfacher ist als eine Grösse Y, falls jemand X verstehen kann, ohne dass er Y versteht, aber nicht umgekehrt.

Diese verschiedenen Facetten der Einfachheit sollen hier nicht ausführlich diskutiert werden. Für eine ausführlichere Darstellung mit illustrierenden Beispielen lese man die Seiten 87-90 in Swinburnes Buch. Er diskutiert auch kurz die Frage, welches Gewicht die verschiedenen Facetten haben (p. 92) und räumt ein, dass beim Vergleich von zwei Theorien nicht immer völlig klar ist, welche nun einfacher ist (p. 92). Auf den Seiten 99-102 geht er zudem auf Versuche ein, die fundamentale Rolle der Einfachheit im Zusammenhang der intrinsischen Wahrscheinlichkeit wegzudiskutieren. Insgesamt sind die Bemühungen von Swinburne, die Bestimmung von intrinsischen Wahrscheinlichkeiten auf ein objektives Fundament zu heben, beachtlich und bedenkenswert. Nach meiner Ansicht kann er seine Argumente sehr überzeugend darstellen. Dennoch seien zwei Punkte in seinem Konzept der Einfachheit angesprochen:

1. Es ist nicht nur die Einfachheit, die eine Theorie attraktiv macht, es gibt auch ästhetische Gründe. Besonders in der theoretischen Physik scheinen mir letztere durchaus hohes Gewicht zu haben, obwohl es wiederum nicht einfach ist, diese Art von Ästhetik in Worte zu fassen. Eines der berühmtesten Beispiele einer über alle Massen ästhetische Theorie ist die allgemeine Relativitätstheorie. Ich kenne kaum eine Person, die diese Theorie verstanden hat und sich nicht von ihrem Zauber irgendwie beeindruckt zeigt. Mir scheint, diese Art von Ästhetik ist nicht nur auf Einfachheit zurückzuführen.

2. Swinburne zeigt an vielen Beispielen, wie vernünftig es in der Vergangenheit und heute ist, gleichwertige Theorien anhand ihrer relativen Einfachheit zu beurteilen, aber er vermag nicht, zu erklären, warum eine einfache Theorie der Wahrheit näher kommt. Es ist zwar tatsächlich so, dass der Fortschritt in den Naturwissenschaften u.a. auf dem Prinzip der Einfachheit beruht, und dass es kaum möglich ist, dieses Prinzip aufzugeben, weil dann fast zwangsläufig Chaos folgen würde, aber ich vermisse bei diesen Gründen den Bezug zur Wahrheit. Was macht denn eine einfache Theorie wahrer? Meines Erachtens ist das Postulat der Einfachheit ein Axiom, das zu Beginn vorausgesetzt werden muss und mit welchem die meisten Wissenschaftler intuitiv übereinstimmen, aber das möglicherweise auch anders gewählt werden könnte.

Mir scheint insgesamt, dass trotz der Bemühungen von Swinburne die Objektivität des Bayesianischen Schliessverfahrens nicht wirklich gewährleistet werden kann. Das wurde auch am Beispiel der physikalischen Prior-Wahrscheinlichkeit für die physikalischen $(\theta_1, \theta_2, \ldots)$ Parameter deutlich (Appendix B). Ich wüsste nicht, wie man auf ganz objektive Weise untere und obere Schranken für die Parameter festlegen könnte. Daher kann kaum abgestritten werden, dass dem

BS auch subjektive Komponenten anhaften.[15] Dies wurde besonders im Zusammenhang mit ID deutlich (Kapitel 3), da dort auch weltanschauliche Komponenten in die Argumentation einflossen.

[15]Ich positioniere mich also nahe dem, was Swinburne als „subjective probability“ (p. 70) bezeichnet. Swinburne selber argumentiert für seine „epistemic probability“ (p. 68), wonach es korrekte Kriterien gebe, wie intrinsische Wahrscheinlichkeiten zugewiesen werden sollen, wir diese Kriterien mehr oder weniger kennen würden. Ich bin mir nicht sicher, ob es wirklich solche objektiven, korrekten Kriterien gibt, die uns zwangsläufig der Wahrheit näher bringen.