Die Familie der nicht-deduktiven Schlüsse

Manuel Bremer, Version 4. November 2014

Inhalt

Verwendete Zeichen

Neben den üblichen Formalismen ($\exists$, $\forall$, $\top$, $\supset$, $\varnothing$ …) werden folgend Zeichen verwendet:

$\rightsquigarrow$	kontrafaktisches Konditional („A $\rightsquigarrow$ B“ sagt: „Wenn A der Fall wäre, wäre B der Fall“)
$\Rightarrow$	metasprachliches Konditional in einer Regel
$\curvearrowright$	begründungs- bzw. evidenzbezogenes Konditional („A $\curvearrowright$ B“ sagt: „aufgrund von A kann B als hinreichend begründet angenommen werden“)
$\ni$	Enthaltensein
Prob	(‚Probability’) Wahrscheinlichkeit in einem neutralen Sinne
PROB	epistemische Wahrscheinlichkeit
prob	objektive Wahrscheinlichkeit
prok	kontrafaktische Wahrscheinlichkeit
freq	Frequenz
$\mathfrak{T}$	eine fast sichere (für gewiss gehaltene) Aussage (etwa normalerweise eine Aussage über das eigene Alter oder Geschlecht)
T(A)	A ist wahr (bezüglich eines Satzes A)
F(A)	A ist falsch (bezüglich eines Satzes A)
J(A)	A ist berechtigt annehmbar (‚justifizierbar’)
W(A)	Ein (i.d.R. nicht genauer spezifiziertes) Subjekt weiß, dass A der Fall ist
Ü(A)	Ein (i.d.R. nicht genauer spezifiziertes) Subjekt ist von A überzeugt
M(A)	Ein (i.d.R. nicht genauer spezifiziertes) Subjekt hält A für den Fall (meint, dass A)
$\boxdot$A	naturgesetzlich notwendig, dass A
Z(S)	zukünftig ist es der Fall, dass S [„S“ ist ein nicht-zeitloser Satz]
V(S)	in der Vergangenheit war es der Fall, dass S [„S“ ist ein nicht-zeitloser Satz]
A(S)	jetzt ist S der Fall [„S“ ist ein nicht-zeitloser Satz]
A, B, C …	Zeitlose Sätze (in direktem Gebrauch im Text die betreffenden Sachverhalte/Tatsachen)
S, S_1, S_2…	Sätze, die temporal modifiziert werden müssen, um zeitlos zu sein
F, G, H …	Prädikate (in direktem Gebrauch im Text die betreffenden Eigenschaften)
a, b, c …	Namen (in direktem Gebrauch im Text die betreffenden Individuen)
t, t_1, t_2	Zeitpunkte/-intervalle
Δ, Γ	Mengen von Aussagen

Sprachregelungen

Im Prinzip könnte eine Theorie des nicht-deduktiven Schließens – zumindest zunächst – die gängige Ontologie anderer philosophischen Logiken (also ‚mögliche Welten' etc.) übernehmen. Die vorliegende Untersuchung vermeidet möglichst ontologische Verpflichtungen, die über die Annahme von Repräsentationen und Ausschnitten der raumzeitlichen Wirklichkeit hinausgehen. Dazu sind einige Sprachregelungen nötig bzw. verständlich zu machen.

Behauptung
die in einem Behauptungsakt gemachte Äußerung (ein Sprechvorkommnis).

Aussage
Gehalt einer Behauptung: ein mit Wahrheitsanspruch geäußerter oder gemeinter nicht-indexikalischer Satz (d.h. Verankerung eines in einer Situation geäußerten indexikalischen Satzes durch Ersetzen der indexikalischen Ausdrücke durch Ausdrücke für die situativen Referenten). Logisch äquivalente nicht-indexikalische Sätze sind *gehaltsgleich*. Einer Aussage kann ein Wahrheitswert oder eine Wahrscheinlichkeit zukommen. Aussagen über Wahrscheinlichkeiten kommt ein Wahrheitswert zu.

Prognose
Aussage über die Zukunft [vgl. Abschnitt 3].

Tatsache
Referenz einer wahren Aussage; nur wahre Aussagen haben Referenz. [Vergleiche auch die Bemerkungen in Abschnitt 2 zur ontischen Unterscheidung von Tatsachen und Ereignissen.]

Sachverhalt
Beschreibungs*inhalt* einer Aussage: wir verstehen den Inhalt, indem wir die Aussage verstehen, d.h. wir wissen, auf welche Tatsache die Aussage sich bezöge, wenn sie wahr wäre, wobei dieser Konjunktiv den Realis (dass sie sich auf die Tatsache bezieht, wenn sie wahr ist) einschließt. Einen Inhalt verstehen bzw. sich *auf einen Sachverhalt zu beziehen* heißt nicht mehr, als die Repräsentation (der Aussage) in einer bestimmten Weise zu verarbeiten:

die Inhaltsredeweise muss man nicht ontologisch aufladen. Logisch äquivalente Aussagen sind inhaltsgleich (s.o.).

Prinzip der Unteren Grenze
In einer Kette von Gründen kann die Gesamtstärke der Begründung nicht stärker sein als die Stärke des schwächsten Gliedes der Kette, sie muss jedoch nicht wesentlich schwächer sein.

Individuelle Wahrscheinlichkeit
Wahrscheinlichkeit für das Bestehen einer Tatsache.

Proportion
Naturgesetzlich nicht-kontingente Frequenz (d.h. eine Frequenz, welcher ein nicht-deterministisches Naturgesetz korrespondiert, oder die aus einem solchen folgt).

Propensität
Eigenschaft, für die nicht-deterministische Naturgesetze gelten.

Eigenschaft
Eine (konkrete) Beschaffenheit eines Gegenstandes, und nur insofern abstrakt (im Sinne von ‚nicht selbstständig existierend'). Mehrere Gegenstände können ähnlich/gleich beschaffen sein. Mit ‚Eigenschaft' wird dann auch diese ähnliche/gleiche Beschaffenheit im Allgemeinen gemeint (d.h. es wird distributiv generalisiert). (Analog sind Relationen zu verstehen.)

Determinante
Eigenschaft, für die deterministische Naturgesetze gelten.

Wahrscheinlichkeitseigenschaft
Eigenschaft, auf die ein genereller Term referiert, in dem probabilistische Ausdrücke vorkommen (z.B. PROB(F) = r – die Eigenschaft mit der subjektiven Wahrscheinlichkeit r für F gehalten zu werden). Genaugenommen handelt es sich also um eine Relation (meist zu einem kognitiven Subjekt).

Einleitung

‚Nicht-deduktives Schließen' kann als Sammeltitel für alle Formen des Schließens verstanden werden, die nicht deduktiv im Sinne einer Möglichkeit der Formalisierung im Rahmen eines korrekten, wahrheitsvererbenden, monotonen und transitiven formalen Ableitungssystem sind.

Dies umfasst so unterschiedliche Weisen des Schließens wie:

a. Schlüsse auf die beste Erklärung (Abduktion)
b. Nicht-monotones Schließen (z.B. Default-Logiken)
c. Statistisches Schließen und Schätzen
d. Enumerative und direkte Induktion
e. Wahrscheinlichkeitsschließen
f. Baysianistische Bestätigungstheorien.

Oft wird mit ‚nicht-deduktiven Schließen' die Idee eines nicht zwingenden (also eines obwohl im Sinne des Kanons gezogenen, aber dennoch enttäuschbaren) Schließens verbunden. Daran anknüpfend kann man versuchen, eine allgemeine Systematisierung und Theorie solchen nicht-deduktiven Schließens zu entwickeln. Die Grundidee dabei besagt:

> (NDS1) Verschiedene Formen nicht-deduktiven Schließens (Abduktion, nicht-monotones Schließen und verschiedene Formen der Induktion) lassen sich als *enthymisches Schließen* auffassen.

Geschlossen wird auf eine Weise, die bei Hinzufügen einer in den Daten nicht gegebenen Prämisse (der Systematizitätserhöhung, Abwesenheit von Ausnahmen, Gleichförmigkeit) die Konsequenz deduktiv abzuleiten gestattete. Das nicht-deduktive Argument ist nicht das Argument, das durch die Hinzufügung der fehlenden Prämisse entsteht, denn dieses wäre (in der Regel) ein deduktives Argument. Das nicht-deduktive Argument bezieht seine Plausibilität aus unserer Bereitschaft, die fehlende Prämisse *bis auf Weiteres* einzuräumen. Die Kehrseite dieses Einräumens liegt im Einräumen der Enttäuschbarkeit dieser Unterstellung (der tatsächlichen Falschheit einer als wahr eingeräumten Prämisse). Nicht-deduktives Schließen lässt sich somit durch eine besondere *Einstellung*, die Risikobereitschaft und Enttäuschbarkeit einschließt, charakterisieren – nicht einfach durch eine besondere ‚Logik' im Sinne eines speziellen formalen Systems.

Das im enthymischen Schließen anvisierte vollständige Argument lässt sich im Rahmen der deduktiven Logik (oft über Disjunktionsbeseitigung oder Universelle Spezialisierung bezüglich der fehlenden Prämisse) rekonstruieren. Hier besteht kein neuer Erörterungsbedarf. Zu erörtern sind indessen Regeln, wie mit der Enttäuschbarkeit der gemachten Unterstellung umzugehen ist. Hier stellen sich neue Fragen der philosophischen Logik. So mag es beispielsweise iterierte Enttäuschung (d.h. Enttäuschung der eine Unterstellung untergrabenden Gegenevidenz) und diesbezüglich neue Ableitungsbegriffe (wie relative Ableitbarkeit und finale Ableitbarkeit) geben. Ebenfalls muss eine Systematik des nicht-deduktiven Schließens entwickelt werden (beispielsweise orientiert an der Art der jeweiligen Unterstellung).
Enthymatisches Schließen darf man insofern nicht mit deduktivem Schließen, obwohl es mittels diesen verstanden werden kann, verwechseln: Der Witz des nicht-deduktiven Schließens liegt gerade im *Einräumen seiner Enttäuschbarkeit*, was ein deduktives Schließen, das einfach eine Prämisse unterstellt und nicht erwähnt, gerade nicht einräumt. In der gemachten Unterstellung geht nicht-deduktives Schließen über die Daten *hinaus*. Wer dies verlässlich tut, stellt sich als guter nicht-deduktiver Räsonierer heraus.

Ein Beispiel: Statistische Syllogismen können scheinbar zu ‚induktiven Inkonsistenzen' führen, wie schon Hempel beklagte:

1. Fast alle Düsseldorfer sind modisch.
2. Max ist ein Düsseldorfer.

∴ (Sicherlich) Max ist modisch.

Zugleich jedoch:

1. Fast alle Philosophen sind unmodisch.
2. Max ist ein Philosoph.

∴ (Sicherlich) Max ist unmodisch.

Eine ‚induktive Inkonsistenz' tritt auf, akzeptiert man beide Konklusionen, da man starke induktive Argumente für sie hat.
Um dies zu umgehen wurden zwei Auswege angeboten:

(i) Man verlangt nach ‚totaler Evidenz'. Wahrscheinlichkeitslogisch müssen bei totaler Evidenz die Wahrscheinlichkeiten der beiden Konklusionen komplementär sein, sodass nicht beide Argumente stark sein können, vielleicht müssen sogar beide revidiert werden.

(ii) Man rekurriert auf logische Wahrscheinlichkeiten, welche allein die Stützungsrelation zwischen Prämissen und Konklusion betreffen, insbesondere die bedingte Wahrscheinlichkeit der Konklusion gegeben die Prämissen. Aus dieser Perspektive könnten beide Konklusionen durch die erwähnten Prämissen sehr wahrscheinlich gemacht werden. Es besteht gar kein Konflikt.

Vorschlag (i) ist psychologisch und epistemisch zu anspruchsvoll. Wir haben nie die ‚totale Evidenz', da wir nicht allwissend sind. Vorschlag (ii) verteidigt zwar die hohen bedingten Wahrscheinlichkeiten, löst indessen das Problem nicht, außer wenn zugleich verboten wird, hier per Konditionalisierung (auf die hoch wahrscheinliche) Aussage zu schließen. Im Sinne eines enthymischen Ansatzes zur Induktion lässt sich das Problem anders lösen: Im ersten Schluss wird unterstellt, dass Max ein typischer Düsseldorfer ist und wir keine weiteren relevanten Informationen besitzen. Die zweite Prämisse im zweiten Schluss untergräbt jedoch diese Unterstellung mit einer relevanten, spezifischeren Information, sodass die Konklusion aus Schluss 1 bis auf Weiteres zugunsten der Konklusion aus Schluss 2 zurückzuziehen ist.

In diesem Bild nicht-deduktiven Schließens sind sowohl deduktives als auch nicht-deduktives Schließen die zwei Vorgehensweisen, mit denen wir bestrebt sind, die Systematizität (Kohärenz) unserer Meinungssysteme und Theorien zu erhöhen: Deduktion und Wahrscheinlichkeitsschlüsse entfalten den *Zusammenhang* gegebener Daten *weiter*, nicht-deduktives Schließen stellt in Erweiterung der Daten auf eine nicht beliebige, sondern geregelte Weise neue Zusammenhänge her. Hier findet sich das – kantische – Konzept der Vernunft als Vermögen *aller* Schlussweisen, die das System der Erkenntnisse bilden und entfalten, wieder! Beide Formen des Schließens vereinigen sich in der grundlegenden Funktion der Logik im Allgemeinen.

In einem Vorbegriff lässt sich sagen:

(NDS2) Nicht-deduktive Argumente sind (i) enttäuschungsanfällig (somit nicht deduktiv), aber (ii) teilen mit deduktiven Argumenten die Funktion den Status der Prämissen (in der Regel deren Glaubwürdigkeit oder Chance auf Wahrheit) auf die Konklusion zu vererben, wobei (iii) beim Übergang zur Konklusion dieser Status abgeschwächt wird.

In (NDS1) und (NDS2) ausdrücklich nicht eingeschlossen ist das übliche Wahrscheinlichkeitsschließen. Dieses beruht auf einer eigenen logisch-mathematischen Grundlage und darauf bezogenen Korrekheitsidealen (wie ‚Dutch-book'-Argumenten). Wahrscheinlichkeitsschließen ist daher zwingend, wenn auch mit dem Unterschied, dass Sätzen bzw. Ereignissen Wahrscheinlichkeiten statt Wahrheitswerte zukommen. Wahrscheinlichkeitsschlüsse erzwingen zwar nicht ihre Konklusion, sie machen sie jedoch wahrscheinlicher. Vererbt wird hier nicht die Wahrheit von den Prämissen auf die Konklusion, sondern Wahrscheinlichkeit (und damit indirekt Grade der Glaubwürdigkeit). Als eine philosophisch wichtige Frage stellt sich, wie und ob subjektive von objektiven Wahrscheinlichkeiten zu trennen sind. Die Erforderlichkeit einer Theorie des Wahrscheinlichkeitsschließens kann dabei darin liegen, nicht neben den zwingenden, deduktiven Schlüssen eine andere Art des Schließens zu suchen, sondern anlässlich des objektiven Vorliegens von Wahrscheinlichkeiten in der Wirklichkeit die angemessene Form des Schließens mit Feststellungen bezüglich solcher Wahrscheinlichkeiten zu systematisieren. Für die Theorie der Korrektheit von Wahrscheinlichkeitsschlüssen kann eine Fundierung objektiver Wahrscheinlichkeiten in einer Ontologie und Modelltheorie der Frequenzen und Proportionen den Ansatz bieten.

Die hier vorliegende Erörterung nicht-deduktiven Schließens legt den Schwerpunkt auf die Identifikation und Erläuterung von Schlussprinzipien, denen rationale Akteure in ihrem Räsonieren folgen sollten. Der Fokus richtet sich damit auf Schlussweisen, die uns *bewusst zugänglich* sind und von uns in reflektierender Deliberation – wenn auch nicht in der fachterminologisch exakten Fassung – eingesetzt werden. Ansatzweise können wir sie in einer mindestens umgangssprachlichen Formulierung anführen, wenn wir gegenüber Diskurspartnern nicht-deduktive Schlüsse ausführen oder rechtfertigen. Schulung und methodische Reflexion steigern unsere Fähigkeit der Erläuterung, unsere Fähigkeit, uns diesbezüglich verständlich zu machen. Partiell kann Schulung auch unsere Fähigkeit zu nicht-deduktiven Schließen verbessern, im Allgemeinen werden wir aber immer schon Prinzipien dieser Art befolgt haben. Die philosophische Erörterung solcher Prinzipien bemüht sich, diese Kompetenzen und dieses partielle Verständlichmachen expliziter und systematischer zu fassen. Insofern werden in der philosophischen Erörterung auch Formalismen vorkommen, mit denen wir so im Alltag nicht bewusst umgehen.
Auf der anderen Seite müssen Prinzipien, mit denen wir bewusst umgehen bzw. die uns bewusst zugänglich sind, so sein, dass sie nicht Fertigkeiten und Wissensbestände verlangen,

die uns in psychologisch realistischer Weise nicht zugesprochen werden können. „psychologisch realistisch“ meint in diesem Kontext, dass die Abwesenheit von bestimmten Repräsentationen in der Introspektion (etwa die von reel-wertigen Wahrscheinlichkeitseinschätzungen) auch dagegen spricht, Wissen über solche Werte als wesentlichen Bestandteil von Schlussprinzipien anzusetzen. Begrenzungen unserer Leistungsfähigkeit als kognitiver Systeme (etwa in der Menge betrachtbarer Daten und bezüglich des Zugangs zu im Prinzip herleitbaren Daten) beschränken die in solchen Schlussprinzipien sinnvoll annehmbaren Idealisierungen. Damit unterscheiden sich die hier betrachteten Schlussprinzipien von denen, die in Theorien maschinellen Lernens oder baysianischer Netzwerke angenommen oder untersucht werden. Es mag sein, dass in unserer Kognition Module eingebaut vorliegen, die sich angemessen als Wahrscheinlichkeitsnetzwerke beschreiben lassen. Im Unterschied zu Prozessen der Deliberation sind uns die Vorgänge in solchen Modulen allerdings nicht bewusst zugänglich. Nicht-deduktive Deliberation orientiert sich an Schlussprinzipien, die bei Bedarf (etwa dem Ausführen eines Begründungsganges) anführbar sein müssen.
Die hier entwickelte philosophische Erörterung des nicht-deduktiven Schließens orientiert sich daher weder an mathematischen Idealisierungen (wie dem validen Berechnen von numerischen Bestätigungsgraden) noch am Ideal korrekter und deduktiv vollständiger formaler Systeme. Die Erörterungen orientieren sich an dem Ideal ein – eventuell noch unvollständiges – System von Grundsätzen nicht-deduktiven Schließens in seinem systematischen Zusammenhang (also über deren bloße Kollektion hinausgehend) zu entfalten und in einem einheitlichen Format und vor dem Hintergrund eines geteilten ontologischen und erkenntnistheoretischen Horizontes darzustellen.

Der Fokus liegt auf Schluss- und nicht auf Evidenzprinzipien. *Schluss*prinzipien betreffen erlaubte und untersagte Übergänge von Prämissen zu Konklusionen. *Evidenz*prinzipien betreffen das Vertrauen in bestimmte Quellen von Meinungen. Dazu gehören beispielsweise:

(E1) Vertraue (bis auf Weiteres) deinen Wahrnehmungsurteilen.

(E2) Vertraue (bis auf Weiteres) deinen introspektiven Urteilen bezüglich deiner mentalen Zustände.

Solche Evidenzprinzipien fundieren oft die *Annehmbarkeit von Prämissen*.

Mit der Theorie der Wahrscheinlichkeit verbinden sich auch eine Reihe schwieriger ontologischer und erkenntnistheoretischer Fragen, die in Einführungstexten zum Wahrscheinlichkeitsschließen in der Regel vernachlässigt werden. Oft wird gleichermaßen und je nach intendierter Anwendung scheinbar beliebig zwischen subjektiven und nicht-subjektiven Verständnissen von ‚Wahrscheinlichkeit' hin und her gegangen. Gelegentlich wird eine Auffassung von Wahrscheinlichkeit ausschließlich benutzt. Demgegenüber soll hier sowohl subjektiven Wahrscheinlichkeitseinschätzungen als auch objektiven Wahrscheinlichkeiten ein Platz eingeräumt werden, ohne die beiden zu vermischen. Vielmehr kommt es oft darauf an, zu unterscheiden, über welche Sorte von Wahrscheinlichkeiten gesprochen wird. Hier ergeben sich ganz verschiedene Schlussweisen bzw. kann die Korrektheit einer Schlussweise davon abhängen, in welchem Sinne die beteiligten Wahrscheinlichkeiten verstanden werden.
Ausflüge in die Ontologie und Erkenntnistheorie ergeben sich daher zwangsläufig.
Die Erörterungen bewegen sich dabei erkenntnistheoretisch im Rahmen eines realistisch verstandenen Internalismus.[1] / [2]

---

1 Vgl. zum allgemeinen Ansatz Bremer, *Philosophische Semantik*.

2 Für die Erörterungen hier waren als Quellen besonders wichtig die Arbeiten von John Pollock: *Nomic Probability and the Foundations of Induction*, bezüglich der Theorie objektiver Wahrscheinlichkeit sowie *Cognitive Carpentry* und *Thinking About Acting* bezüglich der Prinzipien der Deliberation, Reschers Buch *Induktion* hinsichtlich der Idee enthymischen Schließens, Lenzens *Glauben, Wissen und Wahrscheinlichkeit*, an das der zweite Abschnitt anknüpft. Zum Wahrscheinlichkeitskalkül, wie er in Abschnitt 3 angerissen wird, gibt es eine größere Menge von zugleich erörternden Einführungen (wie Skyrms *Einführung in die Induktive Logik* oder Hackings *An Introduction to Probability and Inductive Logic*). Ähnliches gilt für das nicht-monotone Schließen, wie es in Abschnitt 7 angerissen wird (siehe etwa Brewkas *Nonmonotonic Reasoning* oder Besnards *Introduction to Default Logic*). Die Erörterungen hier verfolgen im Horizont dieser Quellen das *systematische* Anliegen, zumindest im Ansatz die Prinzipien nicht-deduktiven Schließens in einer zusammenhängenden Konzeption zu erläutern. Die angegebene Literatur hat dazu die entscheidenden Bausteine geliefert, die vorliegenden Erörterungen sind aber keine Darstellung und Diskussion dieser Positionen und stellen sich nicht in die Gefolgschaft eines der Autoren. Es lassen sich deren Theorien teilweise schwerwiegend kritisieren, wie dies auch in der Sekundärliteratur geschehen ist. Entsprechende zumindest z.T. historische Auseinandersetzungen werden hier nicht verfolgt und sollten daher auch nicht in der Kritik der Erörterungen hier mit einer systematischen Kritik verwechselt werden.

2. Komparative Wahrscheinlichkeit

In vielen Kontexten, in denen keine statistische Informationen vorliegen, beruhen Wahrscheinlichkeitsbehauptungen auf Einschätzungen bzw. Vergleichen bezüglich der Wahrscheinlichkeit des Vorliegens zweier Sachverhalte (bzw. der Wahrheit zweier Aussagen).

Es handelt sich hier um *epistemische* Wahrscheinlichkeiten (*subjektive* Wahrscheinlichkeiten), vor allem um Wahrscheinlichkeitseinschätzungen, dass eine Aussage wahr ist.

Insofern sich solche subjektiven Einschätzungen Repräsentationen bedienen, können als Objekte ihrer Beurteilung Aussagen genommen werden. Die Beurteilung bezieht sich auf die Wahrheit einer Aussage bzw. das Vorliegen der Referenz der Aussage (d.h. die Wahrscheinlichkeit, dass sich die Aussage auf eine Tatsache bezieht, dass diese Tatsache vorliegt). Eine Unterscheidung von Tatsachen (im engen Sinne als Zuständen) und Ereignissen (als Übergängen zwischen Zuständen) ist hierbei oft nicht nötig: man mag Ereignisse einteilen in statische (d.h. Tatsachen) und dynamische (Ereignisse im engeren Sinne), oder Tatsachen einteilen in zeitlich ausgestreckte (d.h. Ereignisse) und zeitlich nicht-ausgestreckte (Tatsachen im engeren Sinne). Wir kommen auf diese ontischen Differenzen nur zurück, wenn sie einen wichtigen Unterschied machen. Das Gleiche gilt für die Unterscheidung von Aussage und Referenz der Aussage.

[Deshalb wird zumeist ein einheitlicher Symbolismus gebraucht: die Sätze „A“, „B“ etc. beziehen sich im Falle ihrer Wahrheit auf die Tatsachen oder Ereignisse A, B etc.]

Als Grundrelation im Bereich komparativer Wahrscheinlichkeit soll dienen: Kw(A,B).
„Kw(A,B)“ besagt: „A ist komparativ höchstens so wahrscheinlich wie B“. Dies gilt für ein Subjekt, das zu unterscheiden in der Regel nicht nötig ist, solange nicht Prinzipien intersubjektiver Epistemologie oder intersubjektiven Schließens untersucht werden.
„Kw(A,B)“ ist wahr genau dann, wenn $PROB(A) \leq PROB(B)$. Diese Äquivalenz kann als Wahrheitsbedingung fungieren oder als Austauschprinzip im Herleiten. Insofern wir Schließen mit komparativen Wahrscheinlichkeiten als Erweiterung oder Form des Wahrscheinlichkeitsschließens auffassen, können wir „Kw“ auch definitorisch einführen:

(D0) $Kw(A,B) \stackrel{def}{=} PROB(A) \leq PROB(B)$

Nun kann man „A ist wahrscheinlicher als B“ definieren:

(D1) $Kw+(A,B) \overset{\text{def}}{=} Kw(B,A) \wedge \neg Kw(A,B)$

Und „A ist weniger wahrscheinlich als B”:

(D2) $Kw-(A,B) \overset{\text{def}}{=} \neg Kw(B,A)$

Sowie „A und B sind gleichwahrscheinlich“:

(D3) $Gw(A,B) \overset{\text{def}}{=} Kw(B,A) \wedge Kw(A,B)$

Angelehnt an den Wahrscheinlichkeitskalkül sollen für die Relation Kw zunächst folgende Bedingungen gelten:

(B1) $Kw(\bot,A)$

Falsum (ein Widerspruch) ist subjektiv am unwahrscheinlichsten.

(B2) $Kw(A,\top)$

Verum (eine Tautologie) ist am wahrscheinlichsten.

(B3) $Kw+(\top,\bot)$

Es gibt Grade der subjektiven Wahrscheinlichkeit.

(B3’) $Kw+(\mathfrak{T},\neg\mathfrak{T})$

Subjektive Alltagsgewissheiten können die Rolle von Grenzwerten spielen.

(B4) $Kw(A,B) \vee Kw(B,A)$

Konnexität der Vergleichsbeziehung – eine psychologisch problematische Annahme. Drückt man *Ignoranz* als „$Kw(A,B) \wedge Kw(B,A)$“ aus, scheint (B4) allerdings harmlos.

(B5) $Kw(A,B) \wedge Kw(B,C) \supset Kw(A,C)$

Transitivität der Vergleichsbeziehung – eine Anforderung minimaler Konsistenz in den Wahrscheinlichkeitsurteilen.

(R1) $\vdash(A\equiv B) \Rightarrow ((Kw(C,A) \equiv Kw(C,B)) \wedge (Kw(A,C) \equiv Kw(B,C)))$

Logische Äquivalente sollten/müssen sich substituieren lassen.

Insbesondere (R1) ist eine starke Idealisierung, da eine vorhandene logische Äquivalenz einem epistemischen Subjekt ja nicht bekannt sein muss. (R1) sollte also im Vordersatz weiter eingeschränkt werden. Aber um einige allgemeine Behauptungen über ableitbare Theoreme einer Theorie der komparativen Wahrscheinlichkeit zu machen, sei zunächst davon ausgegangen. Die jeweilige Äquivalenz könnte auch ein *Postulat* für den jeweiligen subjektiven Wahrscheinlichkeitsraum sein, also ausdrücklich angeführt werden. (R1) drückt auch aus, dass Wahrscheinlichkeitsurteile unabhängig von der speziellen sprachlichen *Darstellung* eines Sachverhaltes sind (also das epistemische Subjekt sich durch die verschiedenen Darstellungen indifferent auf deren vermeintliche Referenz bezieht). Dies wird

man insbesondere im Rahmen einer Repräsentationalistischen Theorie des Geistes für höchst zweifelhaft halten.
Epistemische Logiken, zu denen auch diese Theorie der komparativen Wahrscheinlichkeit gehört, verfahren zumeist normativ mit starken Idealisierungen (starken Anforderungen und Erwartungen an rationale Räsonierer). Soll es überhaupt Ableitungen interessanter Theoreme geben, lässt sich dies nicht vermeiden. Bei den Idealisierungen kann allerdings zwischen schwachen Anforderungen an die Rationalität und starken Anforderungen an die Rationalität unterschieden werden. (B4) und (B5) etwa werden wir gelegentlich nicht erfüllen. Im Allgemeinen jedoch sollte uns die Transitivität unserer Beurteilungen – soweit wir sie denn überblicken[3] – keine größeren Schwierigkeiten bereiten. (R1) verlangt mit der Gleichbehandlung logischer Äquivalente wesentlich mehr. (R1) grenzt an die Forderungen logischer Allwissenheit: um logische Äquivalente gleich zu behandeln, muss man sie ja erst einmal als solche kennen (oder sogar herleiten können). Eine Schwierigkeit interessanter epistemischer Logiken liegt also in der *Auswahl* und *Begrenzung* von Idealisierungen. Ein Thema, das uns in der ganzen Untersuchung immer wieder begegnen wird.
Die Liste der Bedingungen wird noch erweitert.

Aufgrund von (B4) und den Definitionen (D1) – (D3) gilt:

(T1) $\mathrm{Kw}{+}(A,B) \vee \mathrm{Gw}(A,B) \vee \mathrm{Kw}{-}(A,B)$

Dieses Theorem drückt die grundsätzliche Vergleichbarkeit zweier Sachverhalte aus.
Aus (B5) und den Definitionen (D1) – (D3) ergeben sich Varianten der Transitivität:

(T2) $\mathrm{Gw}(A,B) \wedge \mathrm{Gw}(B,C) \supset \mathrm{Gw}(A,C)$

(T3) $\mathrm{Kw}{+}(A,B) \wedge \mathrm{Kw}{+}(B,C) \supset \mathrm{Kw}{+}(A,C)$

Beweis[4]:

1.<1> $\mathrm{Kw}{+}(A,B) \wedge \mathrm{Kw}{+}(B,C)$	AE
2.<1> $\mathrm{Kw}(B,A) \wedge \neg\mathrm{Kw}(A,B)$	$\wedge$B, D1, 1
3.<1> $\mathrm{Kw}(C,B) \wedge \neg\mathrm{Kw}(B,C)$	$\wedge$B, D1, 1
4.<1> $\mathrm{Kw}(C,A)$	B5, 2, 3
5.<5> $\mathrm{Kw}(A,C)$	AE
6.<1,5> $\mathrm{Kw}(B,C)$	$\wedge$B, B5, 2, 5
7.<1,5> $\perp$	$\wedge$B, $\wedge$E, PC, 6, 3
8.<1> $\neg\mathrm{Kw}(A,C)$	$\neg$E, $\underline{5}$, 7
9.<1> $\mathrm{Kw}{+}(A,C)$	D1, 4, 8

[3] Vgl. die Bemerkungen in der Einleitung zu einer realistischen Theorie unserer kognitiven Kapazitäten.
[4] Der Beweis illustriert auch das hier verwendete Format von Beweisen, das offensichtlich sein sollte.

10.<> $Kw+(A,B) \wedge Kw+(B,C) \supset Kw+(A,C)$ $\supset$E, $\underline{1}$, 9 $\square$

(T4) $Kw-(A,B) \wedge Kw-(B,C) \supset Kw-(A,C)$

Beweis analog zu (T3).

Da wir in der Regel komplexe Sachverhalte beurteilen, werden wir oft die Bestandteile solcher Sachverhalte miteinander vergleichen und daraus zu Beurteilungen der komplexen Sachverhalte kommen. Auch dabei muss eine Form minimaler Konsistenz gefordert werden, die eine Art Wendegesetz bedingt:

(B6) $Kw(A\wedge B,A\wedge C) \supset Kw(A\wedge\neg C,A\wedge\neg B)$

Wenn wir $A\wedge B$ für höchstens so wahrscheinlich halten wir $A\wedge C$, liegt dies offensichtlich an der Komponente B im Vergleich zur Komponente C, so dass wir konsequenterweise $\neg C$ mit weniger Vertrauen betrachten sollten als $\neg B$.

Aus (D1) und (B6) lässt sich herleiten:

(T5) $Kw+(A\wedge B,A\wedge C) \supset Kw+(A\wedge\neg C,A\wedge\neg B)$

Und mittels des Spezialfalles $A \equiv \top$ und Aussagenlogik ergibt sich ein allgemeines Kontrapositionsprinzip:

(T6) $Kw(A,B) \supset Kw(\neg B,\neg A)$

bzw.

(T7) $Kw+(A,B) \supset Kw+(\neg B,\neg A)$

Additivität bringt zum Ausdruck, dass Zusatzannahmen einen vorhergehenden Vergleich nicht betreffen, sofern diese Zusatzannahmen vom vorhergehenden Vergleich unabhängig sind (d.h. nicht eines der Vergleichsrelata wahrscheinlicher oder unwahrscheinlicher machen).

(B6')[5] $(A\wedge C \equiv \bot) \wedge (B\wedge C \equiv \bot) \supset (Kw(A,B) \supset Kw(A\vee C,B\vee C))$

Daraus folgt:

(R1') $\vdash(A\supset B) \Rightarrow Kw(A,B)$

Die Konsequenzen der Wahrheit einer Aussage (des Vorliegens eines Sachverhaltes) muss man für mindestens so wahrscheinlich halten, wie diese selbst.

Beweis:

1.<>	$A \supset B$	AE
2.<>	$Kw(\bot,B\wedge\neg A)$	(B1)
3.<>	$(B\wedge\neg A)\wedge A \equiv \bot$	PC
4.<>	$\bot\wedge A \equiv \bot$	PC

5 (B6) und (B6') sind logisch äquivalent. Der Beweis ist nicht trivial.

5.<> $Kw(A\vee\bot,(B\wedge\neg A)\vee A)$ (B6’), (PC), 2, 3, 4

6.<> $(A\vee\bot\equiv A)\wedge(B\wedge\neg A)\vee A\equiv B\vee A$ PC

7.<> $Kw(A,B\vee A)$ $\wedge$B, R1, 5, 6

8.<> $(A\supset B)\supset(B\vee A\equiv B)$ PC

9.<> $(B\vee A\equiv B)$ $\supset$B, 1, 8

10.<> $Kw(A,B)$ R1, 7, 9 □

Kommentar: In (1) angenommen wird $\vdash(A\supset B)$, nicht einfach $(A\supset B)$. Mit der Annahme $(A\supset B)$ wäre die Formel in (9) von dieser Annahme abhängig (d.h. nicht selbst ein Theorem). Gebraucht würde dann die noch stärkere Regel

(R1+) $A\equiv B\Rightarrow((Kw(C,A)\equiv Kw(C,B))\wedge(Kw(A,C)\equiv Kw(B,C)))$

die ausdrücken würde, dass alle materiell äquivalenten Aussagen A und B (also insbesondere alle wahren Aussagen) für gleichwahrscheinlich gehalten werden. In Wirklichkeit *sind alle materiell äquivalenten Aussagen auch objektiv gleichwahrscheinlich*, nämlich entweder Prob(A)=Prob(B)=1 oder Prob(A)=Prob(B)=0, aber Wahrscheinlichkeitsbeurteilungen machen nur intensional Sinn (d.h. bezüglich antizipierter, zukünftiger Situationen, die nicht durch Aussagen, sondern durch *Prognosen* beschrieben werden, oder als Ausdruck des Nichtwissens, welche aus einer Menge von Situationsbeschreibungen der Wirklichkeit entspricht, s.u.).

(R1’) als Verschärfung von (R1) wird mittels (R1) bewiesen und geht in die Richtung stärkerer Prinzipien des logischen Abschlusses bzw. des logischen Wissens. Da aussagenlogisch gilt: $\vdash A\Rightarrow\vdash\top\supset A$, folgt aus (R1’)

(R1’’) $\vdash A\Rightarrow Kw(\top,A)$

was ausdrückt, dass man eine beliebige Tautologie für maximal wahrscheinlich hält (d.h. insbesondere unabhängig von deren syntaktischer Komplexität) – eine äußerst zweifelhafte Forderung/Idealisierung. Auch dies wirft wieder ein schlechtes Licht auf (R1) als dem gegenüber (B6) oder (B1) strittigeren Element dieser Herleitungen.[6]

Der Ausweg, überhaupt Prinzipien des logischen Abschlusses in eine Logik der komparativen Wahrscheinlichkeit aufzunehmen, könnte darin liegen, diese Prinzipien grundsätzlich zu

[6] In einer Modellierung wie derjenigen Lenzens werden solche Idealisierungen als wenig problematisch betrachtet. Mittels weiterer Postulate kann komparative Wahrscheinlichkeit dann auch metrisiert werden, indem die Wahrscheinlichkeits*urteile* einer subjektiven Wahrscheinlichkeits*verteilung* korrespondieren. Die resultierende Theorie lässt sich als deduktiv vollständig bezüglich einer solchen Semantik subjektiver Wahrscheinlichkeitsverteilungen (zu möglichen Welten) erweisen. – All das spielt hier im oben erläuterten methodischen Ansatz keine Rolle, da die Postulate als zu stark gewählt werden müssen, um noch psychologisch realistisch zu erscheinen. Selbst wenn Deliberation auf kognitive Module subdoxastisch zurückgreift, die mittels metrisierter (partieller) Wahrscheinlichkeitsverteilungen beschrieben werden können, treten in den Prinzipien der Deliberation die entsprechenden numerischen Berechnungen nicht auf.

subjektivieren. Das heißt, auf objektive Prinzipien wie (R1) – (R1’’) zu verzichten, da sie zu starke Anforderungen stellen. Stattdessen (d.h. anstelle der objektiven logischen Herleitbarkeit) gilt es dann auf Meinungen über Herleitbarkeit und Theoremstatus zu rekurrieren. Das Gebäude der Meinungen und ihrer Zusammenhänge wird damit noch subjektiver. An die Stelle tatsächlicher Folgerung tritt *gemeinte* Folgerung. An die Stelle logischer Wahrheiten treten Aussagen, die für logische Wahrheiten gehalten werden. Auf der anderen Seite entspricht ein solches Vorgehen einem *Internalismus* des Nachvollziehens und Verstehens von Begründungen, die Meinende für ihre Meinungen angeben oder über die sie selbst verfügen.

Für eine solche gemeinte logische Folgerung oder Äquivalenz sei der Symbolismus um Indices für epistemische Subjekte ergänzt: $\vdash_a(A \supset B)$ bzw. $\vdash_a(A \equiv B)$.

An die Stelle von (R1) tritt (R1S):

(R1S) $\vdash_a(A \equiv B) \Rightarrow ((Kw(C,A) \equiv Kw(C,B)) \wedge (Kw(A,C) \equiv Kw(B,C)))$

Entsprechend:

(R1’S) $\vdash_a(A \supset B) \Rightarrow Kw(A,B)$

(R1’’S) $\vdash_a A \Rightarrow Kw(\top,A)$ bzw. $\vdash_a A \Rightarrow Kw(\mathfrak{T},A)$

Die beiden letzteren Regeln ergeben sich analog zu ihren objektiven Varianten, wobei in den Herleitungen (R1) durch (R1S) ersetzt wird. Die Annahme (1) im obigen Beweis von (R1’) stellt sich dann als eine subjektive Meinung zum Vorliegen einer Implikation dar. Bisher angeführte Theoreme (etwa der Aussagenlogik) wie (6) im obigen Beweis von (R1’) müssen jetzt als eigene subjektive Annahmen aufgeführt werden. Eine leere Abhängigkeitsnotierung (wie „12.< >“) drückt immer noch aus, dass es sich beim Satz in der betreffenden Zeile um ein Theorem handelt. Der Unterschied zwischen objektiver und subjektiver logischer Wahrheit soll nicht aufgehoben werden. Auch wollen wir gelegentlich objektive und subjektive Herleitungen vergleichen. Hinzutreten muss eine subjektabhängige Abhängigkeitsnotierung (wie „12.<a>“). Diese drückt nun nicht mehr aus, dass in der betreffenden Zeile ein Theorem steht, sondern, dass das, was in dieser Zeile steht, vom betrachteten Subjekt für ein Theorem gehalten wird.

Zusätzlich muss nun angenommen werden – als eine Art Hintergrundannahme – dass die üblichen Ableitungsregeln (wie im obigen Fall *Modus Ponens*) den epistemischen Subjekten

bekannt sind. Dies ist wieder eine Idealisierung, aber gegeben den (bisher) elementaren Charakter dieser Regeln erscheint sie akzeptabel – und in diesem Fall harmloser als (R1).[7]
Was zeigt nun eine Ableitung? Ableitungen werden damit nicht beliebig. Eine Ableitung zeigt, was gemäß den Meinungen und den *prinzipiell* vom epistemischen Subjekt akzeptierten Standards für dieses folgen müsste. Unachtsamkeit und Irrtum mögen bei einem Subjekt das Ableiten beeinträchtigen. Trotzdem und gerade deswegen wird ein rationales epistemisches Subjekt (und auch ein solches mit beschränkter Kenntnis logischer Theoreme) einen ihm mittels einer seine Standards korrekt befolgenden Ableitung nachgewiesenen Fehler einräumen.
Dieser Internalismus der Begründung kann sogar objektiv falsche Annahmen über logische Wahrheiten integrieren. Solche falschen Annahmen (logische Fehler im engeren Sinne) werden im Unterschied zur Abwesenheit von Annahmen bezüglich logischer Wahrheiten (logischer Ignoranz) allerdings oft schnell in Widersprüche führen, die das Subjekt auf einen Fehler hinweisen.
Als Konsequenz dieser Subjektivierung können nur *relative* Aussagen zur Korrektheit und deduktiven Vollständigkeit eines Folgerns gemacht werden. Für jede objektiv korrekte Herleitung droht, dass dem Subjekt eine benutzte logische Wahrheit nicht zur Verfügung steht. Betrachten wir noch einmal die nun subjektivierte Herleitung von (R1S):

1.<a>	$A \supset B$	AE
2.<>	$Kw(\bot, B \wedge \neg A)$	(B1)
3.<a>	$(B \wedge \neg A) \wedge A \equiv \bot$	PC
4.<a>	$\bot \wedge A \equiv \bot$	PC
5.<a>	$Kw(A \vee \bot, (B \wedge \neg A) \vee A)$	(B6’), (PC), 2, 3, 4
6.<a>	$(A \vee \bot \equiv A) \wedge (B \wedge \neg A) \vee A \equiv B \vee A$	PC
7.<a>	$Kw(A, B \vee A)$	$\wedge$B, R1, 5, 6
8.<a>	$(A \supset B) \supset (B \vee A \equiv B)$	PC
9.<a>	$(B \vee A \equiv B)$	$\supset$B, 1, 8
10.<a>	$Kw(A, B)$	R1, 7, 9 $\square$

Eigentlich müssten, verzichtete man auf jedwede Unterstellung logischen Theoremwissens, alle Annahmen von Theorem mit einem „<a>“ markiert werden. Wie oben schon angesprochen, soll hier indessen von einer zumindest partiellen logischen Kompetenz und

[7] Auch diese Harmlosigkeit weist Grenzen auf, wie sich sofort daraus ergibt, dass sich mittels eines deduktiv vollständigen Systems Natürlichen Schließens ja alle logischen Wahrheiten ohne irgendwelche Annahmen herleiten lassen. Die Grenzen betreffen hier wieder die nur vage andeutbaren Beschränkungen der Übersicht, Vorausplanung und Gedächniskapazität.

dem Vorhandensein logischen Wissens ausgegangen werden. Grundlegende Prinzipien bzw. Axiome (wie (B1) und (B6) in diesem Fall) sollten sich bei Reflektion auf den Begriff des Wahrscheinlichkeitsvergleichs aufdrängen oder jedenfalls zur Verfügung stehen und auf Erläuterung hin als ‚offensichtlich wahr' angesehen werden. Deshalb sollte hier kein Unterschied zwischen „< >" und „<a>" gemacht werden. Schwieriger verhält es sich mit irgendwelchen Theoremen, seien es auch solche der Aussagenlogik. Im obigen Beispiel wurden sie zunächst alle mit „<a>" markiert. Wie bei Annahmen vererbt sich diese Markierung, wenn die entsprechenden Zeilen in einer Regelanwendung benutzt werden (dies betrifft hier die Zeilen (5), (7), (9) und (10)). Das logische Wissen bezüglich (3), (4) und dem linken Konjunkt in (6) sollte, gegeben ein Verständnis von *falsum*, unproblematisch sein. (8) verlangt die im Allgemeinen unterstellbare Kenntnis des *Modus Ponens* und evtl. eine Idee von Disjunktionsbeseitigung (Argumentieren über Fallunterscheidung). Das rechte Konjunkt von (6) verlangt die zusätzliche Kenntnis eines Distributionsgesetzes. Dieses Wissen ist recht elementar, mag einigen aber schon Schwierigkeiten bereiten. Damit wäre für solche epistemische Subjekte auch (R1S) *zumindest in dieser Weise* nicht herleitbar. Ob für sie (R1S) herleitbar ist, wird daher zu einem Problem, das sich ohne Kenntnis des gesamten logischen Wissens des betreffenden Subjektes nicht entscheiden lässt! Unabhängig davon mag ein entsprechendes Subjekt (R1S) als ‚einleuchtend' und logisch wahr ansehen.
Herleitungen liefern damit nur Momentaufnahmen der logischen Kompetenz und des bei einem Subjekt eigentlich anzusetzenden Folgerungsvermögens bzw. eine mittels Annahmenzuschreibungen verfahrende Interpretation des subjektiven Räsonierens.

Während also *externalistisch* betrachtet ein *reliables* Modul des Wahrscheinlichkeitsschließens eine Regel wie (R1) implementieren mag, wird im Folgenden für die Deliberation oft von den subjektivierten Regeln oder Prinzipien ausgegangen.

Komparative Wahrscheinlichkeitseinschätzungen beziehen sich bis hier hin nicht auf andere Wahrscheinlichkeitseinschätzungen derselben oder anderer Personen.
Introspektionsprinzipien beziehen sich auf eigene epistemische Beurteilungen.
Räumt man eingeschränkte positive Introspektion ein, sollte als Ausdruck subjektiver Gewissheit ($Gw(\top,A)$ scheint zu stark, insofern es jeden Zweifel ausschließt) gelten:

(B7) $Kw(A,B) \supset Gw(\mathfrak{T},Kw(A,B))$

Die Fähigkeit positiver Introspektion enthält jedoch per Kontraposition ein Sichverlassenkönnen auf diese positive Introspektion:

(T9) $\neg Gw(\mathfrak{T},Kw(A,B)) \supset \neg Kw(A,B)$

Urteile komparativer Wahrscheinlichkeit warden mit $Gw(\mathfrak{T},Kw(A,B))$ sicher erfasst: das „$\mathfrak{T}$“ drückt nicht den Gewissheitsgrad aus, dass A höchstens so wahrscheinlich wie B ist, sondern den Gewissheitsgrad, dass man das Urteil, A sei höchstens so wahrscheinlich wie B, zu machen.

Ohne Prinzipien der negativen Introspektion muss man eigens postulieren:

(B8) $Kw\text{-}(A,B) \supset Gw(\mathfrak{T},Kw\text{-}(A,B))$

Mittels (R1) könnte man jetzt ableiten:

(T10) $Gw(A,B) \supset Gw(\mathfrak{T},Gw(A,B))$

(T11) $Kw{+}(A,B) \supset Gw(\mathfrak{T},Kw{+}(A,B))$

Ohne (R1) erlaubt vielleicht (R1S) einige dieser Theoreme, die ansonsten eigens zu postulieren wären.

Man könnte nun auch nicht-komparative Wahrscheinlichkeitseinschätzungen aus den komparativen zu gewinnen versuchen. Man ist von etwas empirisch Kontingenten *überzeugt*, wenn man es für mindestens so wahrscheinlich hält wie offensichtliche Aussagen:

(D4) $Ü(A) \stackrel{\text{def}}{=} Gw(\mathfrak{T},A)$

oder

(D4’) $Ü(A) \stackrel{\text{def}}{=} Kw(\mathfrak{T},A)$

um logische Wahrheiten mit zu berücksichtigen.

Man ist geneigt, etwas zu *meinen*, wenn man es für wahrscheinlicher als sein Gegenteil hält:

(D5) $M(A) \stackrel{\text{def}}{=} Kw{+}(A,\neg A)$ [bzw. $Kw\text{-}(\neg A,A)$]

Es sollte intuitiv eine Serialitätsbeziehung gelten:

(T12) $Ü(A) \supset M(A)$

Beweis:

1.<>	$Kw(\neg\mathfrak{T},\mathfrak{T})$	(B3’) und (D1)	
2.<2>	$Ü(A)$	AE	
3.<2>	$Kw(\mathfrak{T},A)$	(D3), (D4), 2	[nach (D4): (2) $\equiv Gw(\mathfrak{T},A)$]
4.<2,a>	$Kw(\neg A, \neg\mathfrak{T})$	(T6), 3	[angenommen (T6) ist *a* bekannt]
5.<2,a>	$Kw(\neg A,A)$	(B5), 1,3,4	
6.<6>	$Kw(A, \neg A)$	AE	
7.<2,6,a>	$Kw(\mathfrak{T},\neg\mathfrak{T})$	(B5), 3, 6, 4	
8.<>	$\neg Kw(\mathfrak{T},\neg\mathfrak{T})$	($\wedge$B), (B3’)	
9.<2,a>	$\neg Kw(A, \neg A)$	($\neg$E) $\underline{6}$, 7, 8	

10.<2,a> $M(A)$ (D1), (D5), 5, 9

[nach (D5): (10) $\equiv$ Kw+(A, $\neg$A)]

11.<a> $Ü(A) \supset M(A)$ ($\supset$E), $\underline{2}$, 10 □

Es gilt auch eine Konsistenzbedingung:

(T13) $M(A) \supset \neg M(\neg A)$

Beweis:

1.<1>	$M(A)$	AE
2.<1>	$Kw+(A, \neg A)$	(D5), 1
3.<1>	$\neg Kw(A, \neg A)$	(D1), 2
4.<1>	$Kw(\neg A,A)$	(B4), 3
5.<1>	$\neg$ $Kw+(\neg A,A)$	(D1), 4
6.<1>	$\neg M(\neg A)$	(D5), 5
7.<>	$M(A) \supset \neg M(\neg A)$	($\supset$E), $\underline{1}$, 6 □

Aus (B4) folgt direkt:

(T14) $Kw(A,A)$

Aus (T14) und entsprechendem logischen Wissen, dass $\vdash A \equiv \top$ im Falle $\vdash A$ bzw. $\vdash (A \equiv A \vee \neg A) \equiv (A \equiv \top)$, in (R1S) – oder mit (R1) – ergibt sich:

(AR1S) $\vdash_a A \Rightarrow Kw(\top,A)$

(AR1) $\vdash A \Rightarrow Kw(\top,A)$

d.h. man hält alle logischen Wahrheiten (die man kennt) für maximal wahrscheinlich. Womit auch klar wird, dass es sich bei diesem Für-wahrscheinlich-halten um eine *Disposition* handeln muss.

Mit (B3) gilt dann:

(AR2S) $\vdash_a A \Rightarrow Kw+(A,\bot)$

(AR2) $\vdash A \Rightarrow Kw+(A,\bot)$

Und da gilt

$\vdash (\neg A \equiv \bot)$

falls

$\vdash A$

gelten:

(AR2S') $\vdash_a A \Rightarrow Kw+(A,\neg A)$

(AR2S'') $\vdash_a A \Rightarrow M(A)$

(AR2') $\vdash A \Rightarrow Kw+(A,\neg A)$

(AR2'') $\vdash A \Rightarrow M(A)$

als Formen der logischen Allwissenheit! (AR2S'') ist jedoch nur die harmlose logische Allwissenheit, dass man alle logischen Wahrheiten, die man kennt, auch dispositional meint. Wegen

(T15) $\mathrm{Kw}(\mathfrak{T},\top)$

folgen aus (AR1) bzw. (AR1S) mit (B5) [Transitivität] und (D4'):

(AR2S''') $\vdash_a A \Rightarrow \ddot{U}(A)$

(AR2''') $\vdash A \Rightarrow \ddot{U}(A)$

Wiederum aufgrund von (B5) und (R1') bzw. (R1S') gelten als Formen des deduktiven Abschlusses:

(AR3S) $\vdash_a(A\supset B), \ddot{U}(A) \Rightarrow \ddot{U}(B)$

(AR3) $\vdash(A\supset B), \ddot{U}(A) \Rightarrow \ddot{U}(B)$

Beweis:

1.<1>	$\ddot{U}(A)$	AE
2.<1>	$\mathrm{Kw}(\mathfrak{T},A)$	(D4), 1
3.<>	$A\supset B$	Annahme "$A\supset B$" sei ein Theorem
4.<>	$\mathrm{Kw}(A,B)$	(R1'), 3
5.<1>	$\mathrm{Kw}(\mathfrak{T},B)$	(B5), 2, 4
6.<1>	$\ddot{U}(B)$	(D4), 5 ▯

Es gilt sogar:

(AR3') $\vdash(A\supset B), M(A) \Rightarrow M(B)$

Beweisidee: man nimmt zur Widerlegung an $\mathrm{Kw}(B,\neg B)$; aufgrund (R1'), das $\mathrm{Kw}(A,B)$ liefert, und (B5) gilt dann aufgrund der ersten Annahme: $\mathrm{Kw}(A,\neg B)$, per Kontraposition: $\mathrm{Kw}(B,\neg A)$; damit und wieder mit (B5) gilt aufgrund der ersten und der indirekten Annahme: $\mathrm{Kw}(A,\neg A)$, im Widerspuch zur zweiten Annahme; also: $\neg\, \mathrm{Kw}(B,\neg B)$, d.h. $M(B)$, *q.e.d.*
Mit dem entsprechenden logischen Wissen gilt:

(AR3S') $\vdash_a(A\supset B), M(A) \Rightarrow M(B)$

Aufgrund der auf $\mathfrak{T}$ (und nicht $\top$) bezogenen Definition von „Ü" gilt nicht unbedingt

(*) $\ddot{U}(A\supset B) \wedge \ddot{U}(A) \supset \ddot{U}(B)$

da $\mathrm{Prob}(\mathfrak{T}) < \mathrm{Prob}(\top) = 1$, so dass sich multiplikative Abschwächung in Konjunktionen wie (*) ergibt – es sei denn man operiert statt mit einer solchen üblichen Semantik subjektive Wahrscheinlichkeiten mit einem *Prinzip der Unteren Grenze*. Dann gilt sogar mit

(**) $M(A\supset B) \wedge M(A) \supset M(B)$

Ein Abtrennungsprinzip, dass in vielen epistemischen Logiken nicht gilt.

Aus den Iterationstheoremen positiver Introspektion ergeben sich aus den Definitionen (D4) und (D5) entsprechende Theoreme positiver Introspektion:

(T16) $Ü(A) \supset Ü(Ü(A))$

(T17) $M(A) \supset Ü(M(A))$

mit (T12) also auch

(T18) $M(A) \supset M(M(A))$

Fügen wir eine minimale Definition des Wissens hinzu:

(D6) $W(A) \overset{def}{\equiv} M(A) \wedge A$

Dann gilt aufgrund von (T12):

(T19) $Ü(A) \wedge A \supset W(A)$

Aufgrund von (D6) und (AR2) gilt:

(AR2'''') $\vdash A \Rightarrow W(A)$

Wegen der Faktizitätsbedingung in der Definition des Wissens (dem zweiten Konjunkt) gilt das entsprechende subjektivierte Prinzip (*AR2S'''') nicht: was ein Subjekt für eine logische Wahrheit hält, muss gar nicht der Fall sein!

Aufgrund der Definitionen (D4) und (D5) gilt – wie auch immer subjektive Wahrscheinlichkeit genauer modelliert wird – aufgrund der basalen Theoreme des Wahrscheintlichkeitskalküls:

(T20) $M(A) \supset PROB(A) > ½$

Wegen der nicht genau angebbaren Wahrscheinlichkeit von $\mathfrak{T}$ lässt sich eine ähnlich interessante Aussagen für „Ü" nicht machen.

Aufgrund von (T20) liegt z.B. folgendes Theorem vor:

(T21) $\neg M(A) \wedge \neg M(\neg B) \supset Kw(A,B)$

Wenn man weder A noch das Gegenteil von B für den Fall hält, kann man A auch für nur höchstens so wahrscheinlich halten wie B.

Beweis:

1.<1>	$\neg M(A)$	AE
2.<1>	$PROB(A) \leq ½$	(T20), (WK)[8], 1
3.<3>	$\neg M(\neg B)$	AE

[8] „(WK)" soll allgemein eine Verwendung von (elementaren) Rechenschritten des Wahrscheinlichkeitskalküls abkürzen, „(Z)" eine Verwendung von (elementarer) Arithmetik. Es gelten hier die entsprechenden Bemerkungen wie zu „(PC)".

4.<3>	$PROB(B) \geq ½$	(T20), (WK), 3
5.<1,3>	$PROB(B) \geq PROB(A)$	(WK), 2, 4
6.<1,3>	$Kw(A,B)$	(D0), 5
7.<>	$\neg M(A) \wedge \neg M(\neg B) \supset Kw(A,B)$	($\supset$E), $\underline{1}$, $\underline{3}$, 6 □

Damit liegt ein Theorem vor, das von der klassifikatorischen epistemischen Beschreibung wieder zurückführt zur komparativen.

Komparative Wahrscheinlichkeiten, sofern sie den Bedingungen (B1) usw. folgen, lassen sich in subjektiven Wahrscheinlichkeitsfunktionen fundieren – bzw. können die komparativen Wahrscheinlichkeitsurteile aufgefasst werden als superveniente bewusste Einschätzungen, die auf subdoxastischen metrisierten Wahrscheinlichkeitszuweisungen beruhen.
Die Berufung auf den Wahrscheinlichkeitskalkül besitzt hier eine explanative Funktion: ein System von Zuständen, die Umstände mit metrisierten Wahrscheinlichkeitsgewichten versehen, kann auf einer tieferen kognitiven Ebene Vorgänge des bewussten subjektiven Einschätzens und Vergleichens implementieren und damit indirekt die kategorischen doxastischen Zustände fundieren. Eine Semantik solcher Zustände mittels subjektiver Wahrscheinlichkeitsfunktionen bietet sich daher als Semantik einer 3ten Person Perspektive an. Aus der Perspektive der 1ten Person können wir die Wahrheitsbedingungen einer solchen Semantik nur im Rahmen komparativer Abschätzungen nachvollziehen oder auf ihr Vorliegen überprüfen. Wir haben keinen direkten Zugriff auf unsere subjektive Wahrscheinlichkeitsfunktion. Aus der 3ten Person Perspektive lässt sich prinzipiell eine subjektive Wahrscheinlichkeitsfunktion zuschreiben, auch wenn selbst hier exakte Werte in der Regel nicht vorliegen werden. Insofern kann die Beschäftigung mit entsprechenden Wahrscheinlichkeitsmodellen und –berechnungen indirekt erkenntnistheoretisch relevant sein.

Komparative Wahrscheinlichkeitsurteile begegnen uns auch in Form des Auftretens von *prima facie* Schlüssen.
Wenn bei Abwesenheit einer recht unwahrscheinlichen Bedingungen $\neg B$ von A auf C geschlossen werden kann, so bleibt im Rahmen der deduktiven Logik doch nur die Darstellung „$A \wedge B \supset C$“. Es muss ausdrücklich B behauptet oder festgestellt werden. Falls wir jedoch B als wesentlich wahrscheinlicher als $\neg B$ einschätzen, unterstellen wir B bis auf Weiteres und verfahren vermittels des *prima facie* akzeptablen „$A \supset C$“.

3. Wahrscheinlichkeit im Allgemeinen

Die allgemeinen strukturellen Bestimmungen der Wahrscheinlichkeit bzw. eines Wahrscheinlichkeitsmaßes gelten gleichermaßen für subjektive und objektive Wahrscheinlichkeiten.

Sie gelten in analoger Weise auch für Frequenzen und Proportionen, auf die objektive Wahrscheinlichkeiten hier zurückgeführt werden.

Nicht alle Theoreme und Prinzipien gelten für kontrafaktische Wahrscheinlichkeiten (s.u.).

Während objektive Wahrscheinlichkeiten (i.d.R.) auf (idealisierte) objektive Verteilungen zurückgehen, drücken subjektive Wahrscheinlichkeiten subjektive Einschätzungen aus, was eher als etwas anderes der Fall ist. Die Bedingungen der Wahrscheinlichkeitsschlüsse fungieren hier als idealisierte Konsistenzbedingungen für eine entsprechende subjektive Wahrscheinlichkeitsfunktion (bezüglich von Aussagen oder Umständen).

Für eine Theorie nicht-deduktiven Schließens kommt es immer wieder darauf an, verschiedene Wahrscheinlichkeitsbegriffe von einander zu trennen:

- subjektive *versus* objektive Wahrscheinlichkeit
- faktische *versus* kontafaktische Wahrscheinlichkeit
- definite *versus* indefinite Wahrscheinlichkeit

Eine definite Wahrscheinlichkeit betrifft eine individuierende Aussage (d.h. eine Aussage mit einer Individuenkonstante) bzw. Prognose.

Und damit zusammenhängend:

- Prognosen *versus* Aussagen
- Proportionen *versus* Häufigkeiten/Frequenzen
- Aussagen über Propensitäten *versus* Prognosen.

Aussagen sind wahr oder nicht-wahr (da falsch oder unbestimmt etc.). Prognosen wie

(1) Morgen wird es regnen.

sind – sofern die Zukunft noch nicht feststeht – weder wahr noch nicht-wahr, sondern das, *was* prognostiziert wird, hat eine bestimmte Wahrscheinlichkeit des Eintretens. Man könnte hier ‚einfache Prognosen' wie (1) unterscheiden von Prognosen, die Aussagen *über* gegenwärtig vorliegende Wahrscheinlichkeiten sind.

Sagen wir

(1') Morgen wird es eher regnen als dass nicht.

so lässt sich dies ausbuchstabieren zu:

(1'') Es besteht (heute) eine Wahrscheinlichkeit x > ½ , dass die Aussage „Es regnet" morgen wahr ist.

Dies drückt entweder eine subjektive Einschätzung aus (bezüglich der Mehrheit der als möglich betrachteten weiteren Weltgeschichten) oder eine objektive Wahrscheinlichkeit (bezüglich der betreffenden Frequenz oder Proportion der mit den bestehenden Fakten und Gesetzmäßigkeiten kompatiblen weiteren Weltgeschichten). Entweder ist (1'') wahr oder nicht wahr. Während der morgige Tag nicht determiniert vorliegt, mag dies für eine Wahrscheinlichkeit, die ihn betrifft, durchaus gelten.

Für ‚einfache' Prognosen ohne eine entsprechende Umformulierung in Aussagen über gegenwärtige Wahrscheinlichkeiten bezüglich des kontrafaktischen Inhaltes der Prognose (d.h. bezüglich des zukünftigen Umstandes, dessen Wahrscheinlichkeit beurteilt wird) gilt nicht das *tertium non datur*, aber das Wahrscheinlichkeitsprinzip der Komplementarität (die Wahrscheinlichkeiten bezüglich der Prognosen A und ¬A summieren sich zu 1).

Im Folgenden wird „Prognose" im Sinne von ‚einfache Prognose' verwendet und von Aussagen *über* Wahrscheinlichkeiten (insbesondere kontrafaktischen Wahrscheinlichkeitsaussagen) unterschieden. Prognosen sprechen nicht über Wahrscheinlichkeiten, sondern ihnen kommen Wahrscheinlichkeiten zu.

Als Ausgangspunkt kann der Begriff der *bedingten* Wahrscheinlichkeit dienen: Prob(A/B) – die Wahrscheinlichkeit von A gegeben, dass B, wobei Prob(B) > 0. Gilt Prob(A/B) = Prob(A/B∨¬B) sind A und B *wahrscheinlichkeitsmäßig unabhängig* von einander, indem das Vorliegen von B genauso viel oder wenig relevant für das (wahrscheinliche) Vorliegen von A ist, wie das Vorliegen von ¬B. Ein Umstand ist für A positiv oder negativ *wahrscheinlichkeitsrelevant* genau dann, wenn Prob(A/B) ≠ Prob(A/B∨¬B).

[Man kann bedingte Wahrscheinlichkeit auch als *partielle Implikation* auffassen, was zu einer Theorie der ‚logischen Wahrscheinlichkeit' führt.]

Die nicht bedingte Wahrscheinlichkeit lässt sich definieren als

(D0) $\text{Prob}(A) = \text{Prob}(A/\top)$

und betrifft die allgemeine (‚unabhängige', ‚a priori') Wahrscheinlichkeit von A.

Konjunktive Wahrscheinlichkeit wird definiert als

(D1) $\text{Prob}(A \wedge B/C) \stackrel{\text{def}}{=} \text{Prob}(A/C) * \text{Prob}(B/A \wedge C)$

(D1') $\text{Prob}(A \wedge B) \stackrel{\text{def}}{=} \text{Prob}(A) * \text{Prob}(B/A)$

Grundlegend sind die folgenden Axiome:

(A1) $\text{Prob}(A/B) \geq 0$

(A1’) $\text{Prob}(A) \geq 0$

(A2) $\text{Prob}(\top/B) = 1$

(A2’) $\text{Prob}(\top) = 1$

(A3) $\text{Prob}(C) > 0, \vdash(\neg(A\wedge B)) \Rightarrow \text{Prob}(A\vee B/C) = \text{Prob}(A/C) + \text{Prob}(B/C)$

(A3’) $\vdash(\neg(A\wedge B)) \Rightarrow \text{Prob}(A\vee B) = \text{Prob}(A) + \text{Prob}(B)$

(A4) $\text{Prob}(C) > 0, \text{Prob}(B/C) > 0 \Rightarrow \text{Prob}(A/B \wedge C) = \dfrac{\text{Prob}(A \wedge B/C)}{\text{Prob}(B/C)}$

(A1) und (A2) drücken eine Normierung der Wahrscheinlichkeitsmaße aus, (A3) Additivität. Eigentlich handelt es sich bei (A3) um eine Regel. Ebenfalls treten bei bedingten Wahrscheinlichkeiten oft Vorbedingungen des Typs Prob(A) > 0 auf. (A1’), (A2’), (A3’) sind die üblichen *Kolmogoroff-Axiome* unbedingter Wahrscheinlichkeit.

Als Regel gilt *Substitution von Logischen Äquivalenten*:

(R1) $\vdash A\equiv B \Rightarrow \text{Prob}(A/C) = \text{Prob}(B/C)$

$\vdash A\equiv B \Rightarrow \text{Prob}(C/A) = \text{Prob}(C/B)$

(R1’) $\vdash A\equiv B \Rightarrow \text{Prob}(A) = \text{Prob}(B)$

Aus (A3) und (A2) ergeben sich:

(T1) $\text{Prob}(\neg A) = 1 - \text{Prob}(A)$

(T1’) $\text{Prob}(\bot) = 0$

(T1’’) $\text{Prob}(A) \leq 1$

Letzteres aus (T1), da $\text{Prob}(\neg A) \geq 0$, wegen (A1).

Aus (D1’) ergibt sich sofort:

(T2) $\text{Prob}(B/A) = \dfrac{\text{Prob}(A \wedge B)}{\text{Prob}(A)}$

was bei Ausgang von unbedingten Wahrscheinlichkeiten als Definition der bedingten Wahrscheinlichkeit dient.

Aufgrund dessen und (R1), da im Falle, dass $\vdash A\supset B$, gilt $A\equiv A\wedge B$, also $\text{Prob}(A) = \text{Prob}(A\wedge B)$:

(AR1) $\vdash A\supset B \Rightarrow \text{Prob}(B/A) = 1$

Entsprechend ergibt sich auch:

(AR2) $\text{Prob}(A) > 0, \vdash(\neg(A\wedge B)) \Rightarrow \text{Prob}(B/A) = 0$

da in diesem Fall $\text{Prob}(A\wedge B)=0$. Allgemein gilt:

(T3) $\vdash A\supset B \Rightarrow \text{Prob}(A) \leq \text{Prob}(B)$

Beweis:

1.<> $(A\supset B) \supset(A\equiv B\wedge A)$ PC

2.<>	$A \supset B$	AE (angenommen also “$A \supset B$” ist ein Theorem)
3.<>	$A \equiv B \wedge A$	($\supset$B), 1, 2
4.<>	$Prob(A) = Prob(B \wedge A)$	(R1), ($\supset$B), 3
5.<>	$Prob(A/B) = r \leq 1$	A1, „r“ neu
6.<>	$Prob(B \wedge A) = Prob(B) * r$	(D1), (=B), 5
7.<>	$Prob(B \wedge A) \leq Prob(B)$	(Z), 6, 5
8.<>	$Prob(A) \leq Prob(B)$	(=B), 4, 7 $\square$

[Dies entspricht dem Prinzip (R1’) aus Abschnitt 2.]
Es gilt ein *Prinzip der Totalen Evidenz*. Zunächst gilt:

(T4) $Prob(A) = Prob(A \wedge B) + Prob(A \wedge \neg B)$

denn $A \wedge B$ ist mit $A \wedge \neg B$ inkompatibel, so dass nach (A3’) gilt: $Prob(A \wedge B \vee A \wedge \neg B) = Prob(A \wedge B) + Prob(A \wedge \neg B)$. (R1) liefert dann (T4). Mittels (D1’) ergibt sich dann:

(T4’) $Prob(A) = Prob(B) * Prob(A/B) + Prob(\neg B) * Prob(A/\neg B)$

Damit haben wir die Bausteine für ein wichtiges Instrument des induktiven Schließens: *Bayes Theorem*.

$$\text{(BT)} \quad Prob(H/E) = \frac{Prob(E/H) * Prob(H)}{Prob(E/H) * Prob(H) + Prob(E/\neg H) * Prob(\neg H)}$$

Die bedingte Wahrscheinlichkeit einer Hypothese H gegeben eine Evidenz E (und damit der *Bestätigungsgrad*, den H erfährt, fall E vorliegt) ergibt sich aus der Wahrscheinlichkeit, welche die Evidenz gegeben die Hypothese hat relativ zur Gesamtwahrscheinlichkeit der Evidenz. Rein rechnerisch werden also Hypothesen durch eintretende Vorhersagen, die eigentlich unwahrscheinlich sind (Prob(E) also niedrig ist), stärker bestätigt als durch weniger riskante Vorhersagen. Und je wahrscheinlicher eine Hypothese eine (zukünftige) Evidenz macht, umso mehr wird sie durch deren Vorliegen bestätigt.
Beweisskizze: Aufgrund der Kommutativität der Konjunktion gilt $Prob(E \wedge H) = Prob(H \wedge E)$; mittels (D1’) löst man dann die Konjunktionen in bedingte Wahrscheinlichkeiten auf und erhält: $Prob(H/E) * Prob(E) = Prob(E/H) * Prob(H)$, sodass Division durch Prob(E) und (T4’) nun (BT) liefern, *q.e.d.*
Für den Baysianismus als der exakten, wahrscheinlichkeitstheoretischen Form der Bestätigungstheorie gilt Ähnliches wie [vgl. Abschnitt 2] für den mathematischen Wahrscheinlichkeitskalkül in seiner Beziehung zu subjektiven (komparativen) Wahrscheinlichkeitseinschätzungen: baysianistische Theorien sind explanativ bezüglich der subjektiven Bewertungen von Bestätigungsbeziehungen. Beispielsweise betrachten wir das Eintreten einer an sich unwahrscheinlichen Konsequenz einer Annahme als stärkere

Bestätigung dieser Annahme als das Eintreten einer ohnehin recht wahrscheinlichen Vorhersage. Dies lässt sich anhand von *Bayes Theorem* schnell einsehen.
Baysianismus modelliert subjektive Wahrscheinlichkeiten. Bestätigung besteht nicht zwischen Fakten. Fakten bestehen zusammen oder nicht. Grade spielen hier keine Rolle. Bestätigung meint die Erhöhung (bzw. Verminderung) des Grades des Für-wahr-Haltens eines Satzes gegeben die Wahrheit/Falschheit oder (veränderte) Wahrscheinlichkeit eines anderen Satzes, in der Regel die Veränderung unseres Zutrauens in eine Hypothese relativ zu (neuen) Daten. Der Grad der Stützung Prob(E/H) lässt sich oft in Ignoranz des Vorliegens eines Faktums oder bezüglich zukünftig auftretender Umstände einschätzen. Die Hauptarbeit beim Bestätigen einer Hypothese gegenüber konkurrierenden Hypothesen besteht im Zumessen oder Schätzen solcher bedingter Wahrscheinlichkeiten (sogenannter ‚likelihoods'). Aus ihnen und den Einschätzungen der unbedingten Wahrscheinlichkeiten Prob(H) etc. liefert *Bayes Theorem* dann eine *objektive* Auskunft, wie wir unser *subjektives* Für-wahr-Halten revidieren müssen, wenn die entsprechende Evidenz auftritt.

Mit dem Wahrscheinlichkeitskalkül verbindet sich daher auch eine Theorie guter Argumente bezüglich der Erhöhung der Glaubwürdigkeit einer Konklusion *gegeben* glaubwürdige/wahrscheinliche Prämissen. Ein gutes Argument spiegelt sich wieder in einer hohen bedingten Wahrscheinlichkeit (dass eine Konklusion wahr ist, falls es die Konjunktion der Prämissen ist).

Eine zentrale Schlussform von Wahrscheinlichkeitsschlüssen ist die *Konditionalisierung* im Falle bedingter Wahrscheinlichkeiten:

(KP1)	1.	Prob(A/B) = r
	2.	B
	∴	Prob(A) = r

Wenn die Bedingung vorliegt, muss das Bedingte *nun* die vormalige bedingte Wahrscheinlichkeit als unbedingte Wahrscheinlichkeit besitzen.
Bayes Theorem sagt uns, wie sehr eine eintreffende Vorhersage eine Hypothese stützt, und mit (KP1) nehmen wir unmittelbar diese erhöhte Wahrscheinlichkeit als die Wahrscheinlichkeit der Hypothese an, wenn die Vorhersage eintrifft.
Eine Abschwächung davon liefert *Jeffreys Regel*, die an das *Prinzip der Totalen Evidenz* anknüpft. *Jeffreys Regel* betrifft Fälle, in denen wir nicht sicher wissen, dass das Bedingte vorliegt. Wir können über eine subjektive Einschätzung verfügen, wie sicher es ist, dass nun –

etwa gegeben eine Nachricht im Radio – eine Bedingung tatsächlich vorliegt. Diese Einschätzung muss nicht wahrscheinlichkeitsmäßig kohärent in unser bisheriges Wahrscheinlichkeitsmaß eingefügt werden, sondern benutzt unsere bisherigen Wahrscheinlichkeitseinschätzungen (unser mutmaßlich kohärentes bisheriges Wahrscheinlichkeitsmaß) um eine Form der Konditionalisierung zu erlauben:

(JR)	1.	Prob(A/B) = r
	2.	Prob(A/¬B) = t
	3.	PROB'(B) = s
	∴	PROB'(A) = r * s + t * (1 – s)

Die Konklusion erinnert an das *Prinzip der Totalen Evidenz*, da sie der Form entspricht: PROB'(A) = Prob(A/B) * PROB'(B) + Prob(A/¬B) * PROB'(¬B). Berechnet wird PROB'(A). Das Ergebnis einer Anwendung von *Jeffreys Regel* ist also keine Einschätzung (subjektive Wahrscheinlichkeit) unseres alten Wahrscheinlichkeitsmaßes, sondern eine neue – mutmaßlich temporäre – Einschätzung, gegeben unsichere situative neue Information über eine Bedingung.

(KP1) wird zusätzlich interessant, ergänzt man es mittels Prinzipien des Für-wahr-Haltens:

(KP2)	1.	Prob(A/B) = r
	2.	r > ½
	2.	J(B)
	∴	M(A)

Falls etwas wahrscheinlicher als sein Gegenteil ist, gegeben eine bestimmte Bedingung, und es lässt sich rechtfertigen, dass diese Bedingung vorliegt, dann sollte man es meinen. (KP2) lässt sich als qualitative Variante von *Jeffreys Regel* ansehen.

Während (KP1) wahrscheinlichskeitsmäßig korrekt ist, handelt es sich bei (KP2) um einen riskanten Schluss, da J(B) $\not\equiv$ B. (JR) geht rational vor, gegeben eine subjektive Einschätzung, wie sehr man das Vorliegen von B annimmt. (JR) bedarf allerdings numerischer Angaben. Eine Variante für Überzeugtsein sieht so aus:

(KP2')	1.	Prob(A/B) = r
	2.	r ≈ 1
	2.	J(B)
	∴	Ü(A)

Bei beiden Schlussweisen handelt es sich um Schlüsse ausgehend von den Begründungsressourcen eines Subjektes.

Mit der Idee objektiver Wahrscheinlichkeit verbinden sich eine Reihe ontologischer und epistemologischer Schwierigkeiten. „prob(Fx/Gx)“ bzw. kurz „prob(F/G)“ drückt eine *generische* Wahrscheinlichkeit aus. Frequenzen und Proportionen sind Tatsachen. Sie haben selbst keine Wahrscheinlichkeit $y < 1$. Auch objektive Wahrscheinlichkeiten beziehen sich auf gegebene Umstände. „prob(F/G) = r“ drückt aus (vermittels der Rückführung von Wahrscheinlichkeit auf idealisierte Frequenzen), dass es eine bestimmte Verteilung der F in den G *gibt*. Es handelt sich um die allgemeine (‚unbestimmte', ‚indefinite') Wahrscheinlichkeit, dass *irgendein* F ein G ist bzw. dass G auch F mit sich bringt.
Von jedem einzelnen F gilt: es ist entweder G oder nicht, die entsprechende Tatsache besteht oder nicht, ontisch ist nichts unbestimmt. Der Satz „prob(F/G) = r“ spricht somit nur allgemein – und insofern ‚unbestimmt' – über die Fs, drückt nur die Verteilung aus, die selbst wiederum definit vorliegt oder regelhaft idealisiert wurde.
Man kann die indefinite Wahrscheinlichkeitsaussage nutzen, um zu einer subjektiven Wahrscheinlichkeit zu gelangen. Dass für ein bestimmtes G auch F gegeben ist, darf – gegeben Prinzipien des Direkten Schließens – mit der betreffenden Sicherheit r geschlossen werden. Ähnlich oben (KP2). Oder man schließt auf die entsprechende subjektive Wahrscheinlichkeit, die den Grad r erbt.
Hier wird nicht auf eine *individuelle objektive* Wahrscheinlichkeit geschlossen, sondern auf einen Sachverhalt (dass dieses G auch F ist). Objektive Wahrscheinlichkeiten sind *in der Regel indefinit*, da gegenwärtige objektive Umstände (im Sinne des Vorliegens eines nicht selbst probabilistischen Sachverhaltes, dass a F ist) nicht wahrscheinlich x (mit $0 < x < 1$) sind, sondern vorliegen oder nicht. Anders verhält sich bei Aussagen über die objektive Wahrscheinlichkeit zukünftiger Umstände, die in den Grenzen nicht-determinierter Naturgesetze möglich sind. Auch Kontrafaktische Wahrscheinlichkeiten sind unter Umständen individuelle objektive Wahrscheinlichkeiten [s.u. und Abschnitt 5].

„wahrscheinlich“ bezieht sich oft nicht darauf, was proportional der Fall ist, sondern entweder (i) was subjektiv der Fall sein könnte relativ zu einem Raum von subjektiven epistemischen Möglichkeiten (subjektive Wahrscheinlichkeit), oder (ii) was der Fall sein *könnte* relativ zu einem Raum von Möglichkeiten (kontrafaktische Wahrscheinlichkeit). Der einfachste Fall kontrafaktischer Wahrscheinlichkeit sind Aussagen *über* die Wahrscheinlichkeit, dass ein *zukünftiges* G ein F ist. Prognosen *haben* die betreffenden Wahrscheinlichkeiten.

Insofern hier davon ausgegangen wird, dass die Zukunft nicht vorliegt, sondern noch ansteht, haben Behauptungen über die Zukunft oft (falls sie Umstände betreffen, die sich in der Zukunft noch ändern können) kontrafaktischen Charakter.
In beiden Fällen (i) und (ii) wird ein Raum von Möglichkeiten (eine Menge von möglichen Welten) betrachtet, der über das Gegebene (die aktuale Welt, die Wirklichkeit) hinausgeht. Insofern sind entsprechende Wahrscheinlichkeitsaussagen *intensional* (bezogen auf das zugrundeliegende semantische Bild/Modell). Dieser intensionale Charakter von Wahrscheinlichkeitskontexten bedingt auch, dass Regeln wie (R1) von logischer Äquivalenz und nicht materialer Äquivalenz ausgehen müssen. Aus den rein materiellen Beziehungen $A \supset B$ oder $A \equiv B$ lässt sich wahrscheinlichkeitsmäßig nichts ableiten.
Subjektive Wahrscheinlichkeit bezieht sich auf einen Spielraum für möglich gehaltener Welten. Objektive Wahrscheinlichkeit kann – wie oben ausgeführt – meist nur eine nicht definite Wahrscheinlichkeit bezüglich von Verteilungen in gleichartigen (z.B. naturgesetzlich gleichartigen) Welten sein. Definite objektive Wahrscheinlichkeiten betreffen Propensitäten (s.u.). Propensitäten als objektive Eigenschaften erlauben Aussagen über Eigenschaften mit Zulkunftsbezug.

Objektive Wahrscheinlichkeiten liegen in der Wirklichkeit vor. Objektive Wahrscheinlichkeiten jenseits von 0 und 1 entsprechen nicht-deterministischen Zusammenhängen von Eigenschaften (etwa prominent in der Mikrophysik). Propensitäten liegen diesen objektiven Wahrscheinlichkeiten zugrunde.
Objektive Wahrscheinlichkeiten benötigen keine extravagantere Ontologie als strikte Naturgesetze bzw. Naturgesetzaussagen. Sie beruhen auch auf der Unterscheidung zwischen Eigenschaften, die andere Eigenschaften determinieren (hier „Determinanten“ genannt) und solchen, die andere Eigenschaften begünstigen (hier „Propensitäten“ genannt). Insofern Determinanten den Grenzfall der Propensitäten bilden, bedarf es ausschließlich der ontologischen Kategorien der Eigenschaften und entsprechender Tatsachen (höherer Stufe). Objektive Wahrscheinlichkeiten können wie Naturgesetze ontologisch auch ohne intensionale Ontologie auskommen.
Im Falle einer nicht-intensionalen Theorie der Naturgesetze (sogenannter ‚Regularitätstheorien’) unterscheiden wir Naturgesetze von anderen Regularitäten, indem man von sie ausdrückenden Allaussagen (allquantifizierte Konditionale, welche eine Beziehung zwischen dem Haben zweier Eigenschaften ausdrücken) zeigt oder begründet, dass sie eine zentrale Rolle in der Systematisierung empirischen Wissens spielen und dass oberflächliche

Allaussagen aus ihnen hergeleitet werden können. Man testet diese Annahme, indem man sie als Postulate eines Rahmens (etwa ‚physisch möglicher Situationen') festsetzt und diesbezügliche Prognosen und Enttäuschbarkeiten (etwa bezüglich anderer wahrer aber nicht als Naturgesetze angenommener Allaussagen) betrachtet oder praktisch, indem man versucht entsprechende Situationen testweise zu realisieren und sie überprüft. Gesichtspunkte der Systematizität (Kohärenz) helfen so, Naturgesetze zu identifizieren.
Einen so bewährten Status kann man durch bestimmte Modaloperatoren ausdrücken (wie ‚physikalisch notwendig': $\boxdot$A).
So lassen sich dann Alternativsituationen sprachlich ausdrücken und simulieren (als ‚mögliche Welten' im Sinne von maximalen Aussagemengen gemäß eines fixierten Standards, hier: physische Möglichkeit/Konsistenz). Es ergibt sich so eine deflationäre Auffassung der Modalontologie. Eigenschaften werden Bestandteile der Semantik solcher Gesetzesaussagen sein, insofern ein Naturgesetz eine Beziehung zweier Eigenschaften *ist*. Mögliche Welten indessen werden nicht als primitive Entitäten erforscht, sondern im Rahmen einer Modellierung als Konstrukte eingesetzt (etwa um die Semantik kontrafaktische Konditionale „A $\rightsquigarrow$ B" zu klären). Eine Klärung der Semantik kontrafaktischer Konditionale folgt derart auf die Klärung von ‚Naturgesetz' und fundiert diese nicht [s.u.].
Man kann die in einer Naturgesetzaussage identifizierte Beziehung zweier Eigenschaften dann ontologisch als besondere Tatsache (etwa als besondere Beziehung zweiter Ordnung zwischen Eigenschaften) auffassen. Eine solche Ontologie von Tatsachen und die entsprechende Ontologie von Eigenschaften muss indessen keinen nicht-extensionalen Charakter besitzen oder primitive Modalentitäten wie mögliche Objekte oder Welten einführen.
Entsprechendes gilt für probabilistische Gesetze. Die entsprechende Beziehung zwischen (den Extensionen von) zwei Eigenschaften F und G betrifft die Frequenz der F in den G. Gesetzesartige Frequenzen werden hier „Proportionen" genannt. Der Frequenz der F in den G entspricht die *objektive Wahrscheinlichkeit*, dass ein G ein F ist (prob(F/G)), wenn es sich bei dieser Beziehung um ein Naturgesetz handelt bzw. diese Proportionalität aus den Naturgesetzen hergeleitet werden kann. Andere Frequenzen mögen *rein* kontingent sein, sofern nicht alles in der Wirklichkeit durch Naturgesetze – seien diese auch probabilistisch – bestimmt wird. Objektive Wahrscheinlichkeiten betreffen – oft idealisierte – Proportionen, welche aus nicht-deterministischen Zusammenhängen zwischen Eigenschaften hervorgehen. Dass es sich bei „prob(F/G) = r" um eine Naturgesetzaussage handelt, wird auf analoge Weise begründet wie im Falle nicht-probabilistischer Naturgesetzaussagen. Hier spielt in der Regel

Induktion eine Rolle oder statistisches Schließen, also selbst schon nicht-deduktive Schlussweisen oder Verfahren (wie das Schätzen).
Eine Eigenschaft, welche nicht-deterministisch eine andere Eigenschaft bedingt, ist eine Propensität des betreffenden Objektes, die dieses physisch besitzt. Propensitäten bezüglich nicht-deterministischer Naturgesetze sind nicht weniger konkret als Eigenschaften, die eine Rolle in deterministischen Gesetzen spielen.
Eine Aussage des Typs „prob(F/G) = r" drückt also ein entsprechendes Naturgesetz aus. Man kann diese Auffassung probabilistischer objektiver Gesetze auch auf die Sozialwissenschaften ausdehnen, falls es dort solche Gesetze gibt.
Bloß kontingente Verteilungen, wie wir sie oft in statistischen Regelmäßigkeiten finden, drücken wir als empirische Frequenzen aus: freg(F/G) = r. Proportionalitäten sind mehr als empirische Frequenzen.
Die Beobachtung von Frequenzen veranlasst Vermutungen über probabilistische Gesetze, nicht jeder Frequenz entspricht jedoch ein Gesetz, und nicht jedes Gesetz schlägt sich in einer beobachtbaren Frequenz nieder. Insofern sich Wahrscheinlichkeitsaussagen aus anderen Aussagen ableiten lassen, lassen sich auch Wahrscheinlichkeitsaussagen für Fälle machen, die bisher nicht beobachtet wurden oder faktisch noch gar nicht aufgetreten sind [s.u.]. Solche Wahrscheinlichkeitsaussagen, die nicht auf der Beobachtung von Frequenzen beruhen, widersprechen aber nicht der Rolle von Frequenzen als *Entdeckungskontext* von Propensitäten – so wenig wie singuläre Konsequenzen über Nichtbeobachtetes aus deterministischen Naturgesetzesaussagen der Rolle der Beobachtung bei der Entdeckung von Naturgesetzen.
Genauso wie bei Naturgesetzen eine *derivative* Redeweise von ‚möglichen Welten' erlaubt ist, erlauben wir uns hier die Redeweise von ‚möglichen Welten' und darin befindlichen ‚möglichen Objekten', die in bestimmten Verhältnissen die Eigenschaften F, G etc. besitzen. Wenn prob(F/G) = r ein Naturgesetz ist, muss das Verhältnis *physisch möglicher* F zu *physisch möglichen* G etwa r sein. Naturgesetze *beschreiben* ja nicht nur die aktuale Welt (die Wirklichkeit), sondern erfassen Beziehungen, die auch in physischen Alternativen zur Wirklichkeit Bestand hätten. Das unterscheidet sie von kontingenten Generalisierungen. Zur Erinnerung: nur solche Welten sind physisch möglich, in denen die Naturgesetze gelten (d.h. in denen alle probabilistischen und nicht-probabilistischen Naturgesetzaussagen wahr sind).
Bei nicht-determinierten naturgesetzlichen Zusammenhängen (objektiven Wahrscheinlichkeiten) kommt hinzu, dass eine *faktische* Verteilung nicht genau der Proportion entsprechen muss. Man denke an folgende Analogie: obwohl ein fairer Würfel gerade auf die 5 gefallen ist, ist die Wahrscheinlichkeit, dass er (im Allgemeinen) auf die 5

fällt, deswegen nicht 1, sondern bleibt 1/6, da bezogen auf die relevanten sechs Alternativsituationen nur eine Situation eine solche ist, in der die 5 fällt. Entsprechend erfassen probabilistische Gesetzesaussagen nicht die faktische Frequenz der F in den G, sondern die Frequenz der F in den G summiert über alle relevanten (hier: naturgesetzlich kontingenten) Alternativen.
Gegeben diese Parallelität bei der Untersuchung nicht-probabilistischer und probabilistischer Naturgesetze sollte man festhalten: Objektive Wahrscheinlichkeiten lassen sich ebenso empirisch und systematisch erforschen wie Naturgesetze – im Kontrast zur ‚rein logischen Theorie' der Wahrscheinlichkeit.

Ausgehend von Naturgesetzen und entsprechender Modellierung möglicher Welten lassen sich nun *kontrafaktische Wahrscheinlichkeiten* einführen. Diese liegen zwischen subjektiven Wahrscheinlichkeiten, bei deren Modellierung durch einen Raum möglicher Welten, je nach Wissensstand des Subjekts, nicht alle Naturgesetze berücksichtigt werden müssen, und solchen individuellen Wahrscheinlichkeiten, die aus den Naturgesetzen und den Randbedingungen folgen. Kontrafaktische Wahrscheinlichkeiten betreffen meist individuelle Wahrscheinlichkeiten, die sich bei Änderung der Randbedingungen (in alternativen Situationen) ergeben. Betrachtet werden die bezüglich der veränderten Randbedingungen ‚nächsten' möglichen Welten und die bezüglich dieser Menge geltenden Frequenzen, Proportionen und deren Konsequenzen. Die hier auftretenden Wahrscheinlichkeiten sind keine subjektiven Wahrscheinlichkeiten – außer bei Subjekten, die alle Naturgesetze kennen. Sie betreffen die *objektiven* Verhältnisse von Eigenschaftsverteilungen und Propensitäten der betreffenden Objekte. Es lassen sich hier objektive individuelle Wahrscheinlichkeitsaussagen machen, insofern der paradigmatische Einwand gegen individuelle objektive Wahrscheinlichkeiten (dass eine Tatsache entweder besteht oder nicht) in diesem Szenario nicht zutrifft. Es lassen sich beispielsweise singuläre Aussagen über die objektive Wahrscheinlichkeit des *morgigen* Regens machen, vor dem Hintergrund der Verteilungen von Regeln in den ‚nächsten' möglichen Weltgeschichten (d.h. denen, welche die Vergangenheit der aktualen Welt teilen).
Individuelle kontrafaktische Wahrscheinlichkeiten referieren auf Eigenschaften, die ein betreffendes Objekt faktisch besitzt und die sich (z.B. zukünftig) in den entsprechenden Situationen im Haben anderer Eigenschaften niederschlagen. Die Eigenschaft kommt dem Objekt *faktisch*, nicht bloß ‚wahrscheinlich' zu: beispielsweise die Eigenschaft mit 80%

Wahrscheinlichkeit in den nächsten 3 Stunden zu zerfallen, behandelte man es soundso (also gegeben bestimmte *kontrafaktische* Situationen): probk(Z(a)/B(a)) = 0.8.
Aussagen *über* Propensitäten in nicht-kontrafaktischen Kontexten sind hingegen schlicht wahr oder falsch. Das Objekt hat die Propensität oder eben nicht. Diesbezüglich können sowohl definite *Prognosen* gemacht werden („Das Atom wird zerfallen“), die einen entsprechenden Bestätigungsgrad besitzen, als auch Prognosen über fortbestehende Propensitäten (also indefinite objektive Wahrscheinlichkeiten). Im Unterschied zur Prognose „Das Atom wird zerfallen“, der eine objektive Wahrscheinlichkeit zukommt, ist die Aussage „Das Atom wird sicher zerfallen“, die (verdeckt mittels „sicher“) *über* eine Wahrscheinlichkeit spricht, jetzt schon wahr oder falsch, je nach dem Charakter der diesbezüglichen Propensität des Atoms.

Objektive Wahrscheinlichkeiten beruhen also entweder auf Proportionen oder auf Propensitäten, wobei die Proportionen mutmaßlich Effekte der vorliegenden Propensitäten sind.

Bernoulis Theorem besagt, dass sich die beobachtbare Frequenz bei Vergrößerung der Stichprobe der tatsächlichen Wahrscheinlichkeit annähert bzw. man diese Wahrscheinlichkeit als Limes der beobachtbaren Frequenz bestimmen kann. Insofern geht der tatsächliche Wahrscheinlichkeitszusammenhang über die beobachtbaren Frequenzen hinaus. Identifiziert man den bedingten Wahrscheinlichkeitszusammenhang mit den Proportionen in der Gesamtheit, so kann man im Sinne von *Bernoulis Theorem* Proportionen als idealisierte Frequenzen auffassen.
Im Falle einer *unendlichen Gesamtheit* können die objektiven Wahrscheinlichkeiten nur auf den ontisch primären Propensitäten beruhen. Man kann sie nicht ausgehend von Frequenzen bestimmen, da diese bei einer unendlichen Gesamtheit entweder trivial sind (i) oder (ii) ordnungsrelativ in einer Limesbetrachtung bezüglich einer Anordnung der aufgetretenen Fälle (man bedenke etwa die Anordnung von geraden, ungeraden oder Primzahlen, die einmal so erfolgen kann, dass die Primzahlen die Frequenz 1/3 besitzen, einmal so, dass sie die Frequenz 1/7 besitzen – für jede beliebige Frequenz). Trivialität tritt in verschiedenen Formen bezüglich einer unendlichen Gesamtheit G auf:

- jede endliche Teilmenge $A \subset G$ hat die Frequenz 0
- jede unendliche Teilmenge $A \subseteq G$ derselben Kardinalität hat die Frequenz 1

- jede unendliche Teilmenge $A \subseteq G$ einer von $|G|$ verschiedenen Kardinalität hat die Frequenz 0.

Überlegungen zur Erforschung probabilistischer Gesetze werden daher in der Regel von endlichen Szenarien ausgehen.

4. Abduktion, Schlüsse auf die beste Erklärung

Aufgegriffen wird in diesem Abschnitt die Idee aus der Einleitung, dass nicht-deduktives Schließen mittels Unterstellungen operiert und oft auf die Erhöhung der Systematizität zielt. „Schluss auf die beste Erklärung“ (kurz „Abduktion“) kann als Titel für vier recht verschiedene Schlussweisen auftreten:

(i) eine Form iterierter Anwendung des Disjunktiven Syllogismus (des Ausscheidens von Alternativen), die auch auftreten kann als eine Form von Faktorenidentifikation (als Finden relevanter Faktoren für ein zu Erklärendes), bei Unterstellung der Vollständigkeit der Disjunktion.

(ii) eine Form der Systematisierung gleichförmiger Daten durch enumerative Induktion (Schluss auf einen Allsatz bezüglich einer zugrundeliegenden Gesetzmäßigkeit), wobei der Allsatz die *beste* Erklärung liefert, indem er erlaubt, aus eben diesem Allsatz, die Daten als Einzelfälle abzuleiten.

In diesen Fällen wäre Abduktion kein besonderes logisches Prinzip, sondern ein Vorgehen Überlegungen zur Erklärung in deduktive Form zu bringen. Im Falle (ii) wäre die Begründung, warum der Allsatz angenommen wird, die Systematisierung (des Wissens oder der Daten). Wird dabei jedoch überhaupt auf den Allsatz *geschlossen*? Es wird legitimiert, warum man von seiner Wahrheit ausgeht. Insofern dies eine Begründung ist, handelt es sich um einen Schluss.

Die Form, die er besitzt, wäre ungefähr:

(AB1) 1. Systematizität ist verlangt.
2. Die Annahme von A erhöht Systematizität.
3. Nichts anderes erhöht die Systematizität mehr
(oder ist jedenfalls nicht bekannt)
∴ A ist anzunehmen

Dies wäre ein deontisch instrumentelles Räsonieren, ausgehend von (1), bzw. im Falle der Klammer von (3) ein *prima facie* Räsonieren. Abduktion ginge somit zurück auf eine oder beide dieser Weisen des Räsonierens.

(iii) eine Form von Anwendung von *Bayes Theorem*.

In diesem Fall wäre „Abduktion“ nur ein anderer Name für komparatives Bestätigungsräsonieren [vgl. Abschnitt 3].

(iv) eine Form der Ergänzung von Annahmen in einem Meinungssystem, um dessen Systematizität (Kohärenz) zu erhöhen.

In diesem Fall läge der Grund für das Einfügen einer Meinung in das Meinungssystem (bzw. in den Korpus einer Theorie) nicht in den einzelnen Daten, sondern im Befolgen der Prinzipien der Kohärenz. Das Räsonieren folgt dann (AB1), unabhängig davon, ob es sich – wie in (ii) – um Allsätze oder um *einzelne* explanative Annahmen handelt.

Eine Form der Abduktion ist die *Kurvenlegung* bei einer Sammlung von Messdaten. Die Kurve soll ‚am besten' zu den Daten passen. Im *Unterschied* zu enumerativen und statistischen Schließen müssen dabei nicht alle Datenpunkte bewahrt werden. In der Regel liegt die Kurve zwischen vielen der gesammelten Datenpunkte.
Bei der Kurvenlegung spielt Einfachheit (als Gesichtspunkt der Systematizität) die entscheidende Rolle. Durch endlich viele Punkte einer nicht-kontinuierlichen Datensammlung können beliebig viele Kurven gelegt werden. Die ungewöhnlicheren Kurven weisen dabei höhere Polynomgrade auf. Einfachere Kurven drücken nicht nur einfachere Gesetzmäßigkeiten (der Berechnung der Punkte auf den Kurven) aus, sondern haben eingebaut die Annahme von der *Gleichartigkeit* der gesammelten Punkte zu allen Punkten (zwischen den gesammelten Punkten). Dadurch korrespondiert diese Einfachheit dem *Prinzip der Erklärungsstärke*: es wird mehr/besser erklärt, insofern *mehr* Daten unter eine Gesetzmäßigkeit fallen, die *weniger* Faktoren enthält.

5. Statistisches Schließen

Beim *Statistischen Schließen* handelt es sich um ein auf exaktere Daten (üblicherweise Frequenzen) zielende Form von Schätzungen.

Schätzen ist eine alltägliche Prozedur der Hypothesengewinnung, sei es bezüglich von Anzahlen

(1) Es sind mehr Clementinen als Äpfel im Korb.

oder Projektionen (von Bewegungen)

(2) Der Ball geht am Tor vorbei.

Es handelt sich zunächst weniger um ein Schlussverfahren als ein Verfahren, Perzeptionen bzw. Informationen zu Einzeldaten in eine Prognose oder zusammenfassende Meinung zu überführen. Schätzen ist ein – idealerweise verlässliches – Verfahren der *Genese* von Meinungen.

Das Verfahren sollte dennoch verlässlich und in seiner Arbeitsweise stabil sein. Stabilität besteht hier darin, dass

(i) Schätzen bei gleichen Daten zu gleichen Ergebnismeinungen führt: *Uniformität*

(ii) eine Schätzung aufgrund von mehr Daten des bei einer anderen Schätzung benutzten Typs mindestens so gut wie diese Schätzung sein muss: *Konvergenz*.

„Verlässlichkeit" drückt die Tendenz aus, per Schätzung wahre Meinungen oder – in messbaren Fällen – nahe an der Wahrheit liegende Meinungen zu haben.

Die Grundregel beim statistischen Schließen ist der *Schluss von der Stichprobe auf die Gesamtheit*:

(SS1S) Gegeben dass freq(F'/G') = r in der Stichprobe s mit $F' \subseteq F$, $G' \subseteq G$

$\therefore$ PROB(F/G) = r

Geschlossen wird in der subjektiven Variante von der beobachteten Frequenz auf eine subjektive bedingte Wahrscheinlichkeit für alle weiteren Fälle. Man kann für ein beliebiges weiteres Objekt c aus der Gesamtheit „c ist ein F, falls c ein G ist" mit der Wahrscheinlichkeit r glauben, bzw. es einfach meinen, falls r > ½.

(SS1O) Gegeben dass freq(F'/G') = r in der Stichprobe s mit $F' \subseteq F$, $G' \subseteq G$

$\therefore$ prop(F/G) = r

$\therefore$ prob(F/G) = r

Geschlossen wird in der objektivierenden Variante von der beobachteten (faktischen) Frequenz auf einen allgemeinen Zusammenhang, der sich in Proportionalitäten und objektiven Wahrscheinlichkeiten widerspiegelt. [Gemäß Abschnitt 3 handelt es sich hier um eine Idealisierung einer Frequenz zu einer Proportion.]

Offensichtlich ist (SS1O) wesentlich riskanter, da die Konklusion schlichtweg falsch sein mag. Man wird verlässlich so nur schließen, wenn es Indizien gibt, dass es sich um einen natur*gesetzlichen* Zusammenhang handelt.

(SS1) ist enthymatisch bezüglich der Unterstellung, dass die Stichprobe repräsentativ ist. Nimmt man diese Unterstellung als weitere Prämisse hinzu – was im Falle von (SS1O) eine *modale* Behauptung bezüglich einer Gesamtheit möglicher Gs wäre – folgen die Konklusionen *deduktiv*, etwa:

(SS1O') 1. Gegeben dass freq(F'/G') = r in der Stichprobe s mit $F' \subseteq F$, $G' \subseteq G$

2. $\boxdot$(freq(F'/G') = prop(F/G))

$\therefore$ prop(F/G) = r

$\therefore$ prob(F/G) = r

mit dem entsprechenden Definitionen von Proportionen, Naturnotwendigkeit und objektiver Wahrscheinlichkeit [vgl. Abschnitt 3]. Noch simpler wäre (2') freq(F/G) = prop(F/G). Dann handelt es sich bei (SS1O) nur um einen Anwendung der Identitätsbeseitigung. Die epistemische Arbeit läge in der Rechtfertigung der Annahme der Repräsentativität. Man könnte diese Annahme allerdings auch als *Normalitätsunterstellung* auffassen:

(SS1O'') 1. Gegeben dass freq(F'/G') = r in der Stichprobe s mit $F' \subseteq F$, $G' \subseteq G$

2. $\neg$J(freq(F'/G') $\neq$ prop(F/G))

$\therefore$ prop(F/G) = r

$\therefore$ prob(F/G) = r

Solange es nicht begründet wurde, dass die Stichprobe abnormal ist, kann man von Repräsentativität ausgehen. Zusammen mit einem Plausibilitätsprinzip, dass mit $\neg$J($\neg$A) *prima facie* A eine Option ist, lässt sich von (SS1O'') zu (SS1O') mit (2') übergehen. Damit ist (SS1O'') allerdings wieder enttäuschbar und nicht deduktiv gültig.

Auch wenn die Rechtfertigung der Annahme der Repräsentativität mutmaßlich andere Prinzipien des nicht-deduktiven Schließens involviert, kann dennoch das Statistische Schließen gemäß (SS1) in all den Fällen, wo diese Annahme nicht einfach als riskante Unterstellung fungiert, als enthymische Variation deduktiven Schließens verstanden werden.

Der Umstand, dass statistisches Schließen mit zunehmender relativer Größe der Stichprobe verlässlicher wird (d.h. die beobachtete Frequenz in einem immer kleinere Interval um die tatsächliche liegt), lässt sich beweisen, hilft dem Schließenden indessen nur, sofern er weiß, wie groß die Stichprobe relativ zur Gesamtheit ist und ob die relevante Eigenschaft F gleichverteilt in den G vorliegt. Diese Information steht oft nicht zur Verfügung. Bei Prognosen (d.h. einer die Zukunft einschließenden Grundgesamtheit) fehlt dieses Wissen in der Regel.

Man kann entsprechend (SS1), wenn $r > ½$, insbesondere wenn $r \approx 1$, auch auf die Wahrheit davon, dass ein beliebiges individuelles G auch F ist, schließen. In diesem Fall wird das Objekt als ein *typisches* G veranschlagt und ein geringes Risiko, falsch zu liegen, wird vernachlässigt.
Diese Variante des statistischen Schließens übernimmt ein höheres Risiko, denn neben die Enttäuschbarkeit der Normalitätsannahme tritt die Fehlerquote $1 - r$.

(SS2) 1. Gegeben dass $freq(F'/G') = r$ in der Stichprobe s mit $F' \subseteq F$, $G' \subseteq G$
2. $r > ½$
$\therefore$ $M(G(a)) \supset M(F(a))$

Direktes Schließen ist ebenfalls eine Form nicht-deduktiven Schließens, bei dem nicht vom Einzelnen auf die Allgemeinheit, sondern von der Allgemeinheit (einer allgemeinen *statistischen* Gesetzmäßigkeit) auf das Einzelne (ein Faktum) geschlossen wird. Die Charakterisierung nicht-deduktiver Schlüsse als im Gegensatz zu deduktiven Schlüssen vom Einzelnen auf das Allgemeine gehend verkürzt also sowohl das Spektrum der deduktiven wie der nicht-deduktiven Schlüsse. Die Arten unterscheiden sich nicht dermaßen einfach in der logischen Form, die Prämissen und Konklusion quantitativ besitzen, sondern in der Weise der Stützung der Konklusion durch die Prämissen oder durch andere Elemente der logischen Form.

Beim direkten Schließen sind zwei Fälle zu unterscheiden:
(i) wir kennen die Frequenz in der Gesamtheit
(ii) wir kennen die Frequenz in der Stichprobe
Aus den Frequenzen lässt sich bei naturgesetzlichen Zusammenhängen zu einer indefiniten objektiven Wahrscheinlichkeit übergehen. Im Falle (ii) bedarf es dazu eines Zwischenschrittes der enumerativen Induktion [vgl. Abschnitt 6].

Von dieser indefiniten Wahrscheinlichkeit prob(F/G) oder auch PROB(F/G) will man im direkten Schließen auf einen *Einzelfall* schließen. Auch hier sind zwei Fälle zu unterscheiden:

(i) der Einzelfall liegt vor, ist uns jedoch nicht bekannt (es handelt sich also nicht um ein Datum, wir benötigen eine subjektive Wahrscheinlichkeit)

(ii) der Einzelfall liegt (noch) nicht vor.

Der zweite Fall birgt die Hürde, dass individuelle *Fakten* keine objektiven bedingten Wahrscheinlichkeiten haben können: „prob(F(a)/G(a))" macht in dieser Allgemeinheit der logischen Form keinen Sinn, da der Gegenstand a entweder die Eigenschaft F hat oder nicht, das zusätzliche Haben von G spielt dabei keine Rolle. Individuelle Fakten haben überhaupt keine Wahrscheinlichkeiten, sondern liegen vor oder nicht. Entsprechende Aussagen haben keine Wahrscheinlichkeits- sondern Wahrheitswerte. Mit temporaler Modifikation ändert sich die Betrachtung, da *Prognosen* einen Wahrscheinlichkeitswert haben.

(3) prob(Morgen regnet es in Köln/Gestern lag ein Tief über den Niederlanden)

macht Sinn, gegeben das Wetter folgt naturgesetzlichen Zusammenhängen. Hier besteht ein enger Zusammenhang mit kontrafaktischen Wahrscheinlichkeiten.

Ebenso sollten Prognosen die Propensitäten eines Gegenstandes betreffend *jetzt* objektive Wahrscheinlichkeiten besitzen, etwa

(4) prob(Nächste Woche ist das Brot schimmelig)

Dagegen sind die Aussagen

(5) Nächste Woche ist das Brot mit 89% Wahrscheinlichkeit schimmelig.

(6) Nächste Woche ist das Brot sicher schimmelig.

heute einfach wahr oder falsch.

Bezüglich definiter objektiver Wahrscheinlichkeiten kommt es also auf die logische Form (insbesondere das Tempus) der involvierten Aussagen an.

Die einfachsten Formen des direkten Schließens sind damit:

(DSS) PROB(F/G) = r

$\therefore$ PROB(F(a)/G(a)) = r

(DSO) prob(F/G) = r

$\therefore$ prob(Z(F(a))/Z(G(a))) = r

Hält man im betreffenden kontrafaktischen Räsonieren die Naturgesetze konstant, lässt sich in (DSO) auch auf „probk(F(a)/G(a))" schließen.

6. Induktives Schließen und das ‚Induktionsproblem'

Induktion wird im Allgemeinen als generalisierendes Schließen verstanden, das über die Datenbasis hinausgeht. Enumerative Induktion im Allgemeinen schließt von einer beschränkten Datenbasis auf eine Gesamtheit, insbesondere von einer finiten Datenbasis bezüglich einer Korrelation auf ein allquantifiziertes Konditional.

Induktive Verallgemeinerungen können, sofern sie Beispiele statistischen Schließen von einer Stichprobe auf die Allgemeinheit sind [vgl. Abschnitt 4], zu deduktiven Schlüssen umgeformt werden, indem jeweils eine unterdrückte Prämisse der Gleichartigkeit der betrachteten Exemplare relativ zur intendierten Gesamtheit eingefügt wird. Beispielsweise:

(1) 1. Tiger fressen Fleisch.
2. Löwen fressen Fleisch.
[3. Tiger und Löwen sind diesbezüglich repräsentativ für alle Katzen.]
$\therefore$ Katzen fressen Fleisch

„diesbezüglich repräsentativ" muss dann entwickelt werden als:

(2) „Tiger(x)" ist repräsentativ bezüglich „Fleischfresser(x)" gdw.
$\forall x(\text{Tiger}(x) \supset \text{Fleischfresser}(x)) \supset \forall x(\text{Katze}(x) \supset \text{Fleischfresser}(x))$

Der Schluss in (1) wird damit deduktiv trivial.

Induktive Verallgemeinerung können bei Unterstellung einer solchen Gleichförmigkeitsannahme als *prima facie* Argumente verstanden werden: Gleichförmigkeit wird bis auf Weiteres angenommen, so dass der Allsatz *prima facie* gilt. Gleichförmigkeit fungiert damit als *default*-Annahme. Induktive Verallgemeinerung ist so aufgefasst eine Form des nicht-monotonen Schließens [vgl. Abschnitt 7].

In beiden Fällen wird induktive Verallgemeinerung an einer andere bzw. an eine bekannte Form des Schließens angeglichen. Im ersten Fall gäbe es hier nur eine besondere Form, Prämissen eines ansonsten deduktiven Schluss zu ergänzen. Aufgrund der besonderen Rolle die Unterstellungen besitzen, empfiehlt es sich aber wieder, enumerative Induktion als Form eine enthymischen Schließens aufzufassen.
Ebenfalls sieht man wieder die entscheidende Funktion einer Einfachheitsannahme (der Gleichförmigkeit).

Insofern induktives Schließen einer Form des Schließens ist, welche Wissenserweiterung durch Systematizität bei der Integration neuer Daten in ein Meinungssystem anstrebt, lässt sich die Frage nach der Rechtfertigung der Induktion auf die Frage nach der Rechtfertigung eines solchen kohärentistischen Vorgehens reduzieren bzw. in diese einbetten. Damit erübrigt sich ein diesbezügliches *besonderes* ‚Problem der Induktion'. Die Rechtfertigung enthymischen induktiven Schließens ergibt sich im Rahmen einer pragmatischen Metarechtfertigung der Kohärenz.

Fragen der Rechtfertigung verweisen auch auf die Frage nach der Beweislast: Ist eher A oder eher ¬A begründungsbedürftig?
Abgesehen von der grundsätzlichen Frage, ob überhaupt für alles eine Begründung verlangt werden darf oder sollte und ob ein entsprechendes Begründungsprinzip selbst nicht lebensweltlich unbegründet ist, muss auch von Zweifeln und Gegenthesen eine Begründung verlangt werden können, wenn sie diese von Behauptungen und Thesen einfordern. Man kann dies als Prinzip des Vorrangs der Meinungserweiterung ausdrücken (‚credulity'):

(CP) Solange es keinen begründeten Zweifel an A gibt, kann von A ausgegangen werden.

(CP) ähnelt einem Evidenzprinzip. Insofern (CP) allerdings auf Begründungspflichten verweist, kann es auch als Schlussprinzip angesehen werden. Ein Prinzip wie (CP) erbringt erhebliche Leistungen für das Ziel der Kohärenz, da es das Ideal der Meinungsvollständigkeit (entweder A oder ¬A zu meinen) unterstützt. Außerdem ermöglicht es offensichtlich einen ökonomischen Umgang mit Ressourcen der Rechtfertigung.
Im Kontext der enumerativen Induktion sowie der Projektion einer Frequenz in der Stichprobe auf die Gesamtheit im statistischen Schließen lässt sich die gemachtet Normalitäts- oder Uniformitätsunterstellung auf (CP) basieren:

(UP) Gegeben die Abwesenheit widersprechender Daten können die vorliegenden Daten als repräsentativ angesehen werden.

Angelehnt an (CP) hat (UP) eher den Charakter eines Evidenzprinzips. Die Abwesenheit widersprechender Daten darf dabei natürlich nicht auf Faulheit oder Borniertheit beim Datensammeln zurückgehen. Genauso wie die Abwesenheit von begründeten Zweifeln in (CP) nicht Ausdruck eigener Beschränktheit sein sollte. Beide Prinzipien richten sich an kognitive Akteure, welche sich auf dem Pfad der rationalen Optimierung der Kohärenz ihres Meinungssystems befinden.

Damit erübrigt sich zumindest ein ‚Problem der Induktion': *dass* wir induktiv vorgehen rechtfertigt sich als Bestandteil unseres Strebens nach Kohärenz.
Ein weiteres ‚Problem der Induktion' (wie sich die Korrektheit induktiven Schließens zeigen lasse) reduziert sich auf die Frage, *ob* induktives Schließen ein erfolgreiches Vorgehen zur Erhöhung von Kohärenz ist.
Eine Rechtfertigung unseres Strebens nach Kohärenz wiederum muss immer selbst kohärentistische Prinzipien in Anschlag bringen und leistet insofern eine Selbsterläuterung eines Methodenkanons, den sie zugleich in eine pragmatische Legimitation einbettet.[9] Im Falle der Induktion zeigt sich dies am meta-induktiven Charakter solcher pragmatischer Legitimation, insofern diese vom bisherigen Erfolg induktiver Methoden deren weiteren Erfolg prognostiziert. Die kohärentistischen Elemente überwiegen allerdings in einer solchen Rechtfertigung (d.h. der Gesamtvergleich zweier Begründungskonzeptionen bezüglich deren Kohärenz), da sich rein meta-induktiv auch eine Anti-Induktion begründen lässt: die anti-induktive Regel immer gegen den bisherigen Datentrend zu prognostizieren schließt meta-anti-induktiv aus dem Misserfolg der Anti-Induktion auf deren Erfolgträchtigkeit. Allgemeine Kohärenzerwägungen (wie z.B. der Zusammenhang von Methoden mit unseren Begriffen von Wahrheit und Effektivität) erst schließen die Anti-Induktion aus.
Im Sinne einer Anwendung von (CP) auf der Meta-Ebene kann man fragen, welche Gründe es denn gibt, an unseren normalen induktiven Vorgehen zu zweifeln.
Eine gelingende Rechtfertigung der Induktion im hier erläuterten Sinne legitimiert dabei ebenfalls die Unterstellungen, die in das induktive Schließen eingehen, insbesondere eine zumindest partielle Uniformität und Gesetzeshaftigkeit der Natur.

Ein spezielles ‚Problem der Induktion' scheint die Frage nach der Projizierbarkeit von Eigenschaften bzw. nach der ‚Natürlichkeit' von Eigenschaften zu sein. Generelle Terme lassen verschiedenste boolsche Kombinationen zu, in dem Sinne, dass ein genereller Term „F( )" ein *definiens* besitzen kann, in dem beliebige boolsche Kombinationen anderer genereller Terme „G_1( )", „G_2( )" … auftreten können. Dadurch decken generelle Terme nicht im Allgemeinen natürliche Eigenschaften ab (d.h. Eigenschaften, die ausschließlich darauf beruhen, dass der Träger der Eigenschaft auf eine spezielle Weise strukturiert ist). Natürliche Eigenschaften treten in naturgesetzlichen Beziehungen auf. Komplexe Eigenschaften (die einem komplexen *definiens* eines entsprechenden generellen Terms korrespondieren) können sich diesbezüglich deviant verhalten. Die Identifikation, welche Eigenschaften stabil in

[9] Etwa in Rosenbergs *One World and Our Knowledge of it* oder BonJours *The Structure of Empirical Knowledge.*

Gesetzen auftreten können bzw. die Klassifikation einer Kandidateneigenschaft diesbezüglich ist das Problem der Projizierbarkeit.

7. Nicht-monotones Schließen

Nicht-monotones Schließen ist enttäuschbar: zusätzliche Information kann erzwingen, einen Schluss zu revidieren (die Konklusion zurück zu ziehen). Typischerweise tritt dies ein, wenn eine Unterstellung, wie sie oft beim nicht-deduktiven Schließen gemacht wird, sich als falsch herausstellt, etwa wenn eine Normalitätsunterstellung in einem speziellen Fall keine Anwendung findet:

(1) Katzen können normalerweise ihre Krallen einziehen.,
$\therefore$ Diese Katze kann (also) ihre Krallen einziehen.
↓ Aber: Diese Katze ist ein Gepard!

Die Zusatzinformation bezüglich des bezogen auf eine typische Katzeneigenschaft untypischen Geparden setzt die Normalitätsunterstellung außer Kraft.
Eine *Unterstellung* hat dabei einen anderen Status als eine Prämisse oder Annahme:

(i) Wäre die Unterstellung eine Annahme, erwiese sich mit der Zusatzinformation die Datenbasis als *inkonsistent*. Da alle Annahmen zunächst gleiche Dignität besitzen, bedarf es dann besonderer Methoden (etwa der Kontraktion), um wieder Konsistenz herzustellen. Auch müssen solche Informationen keinen eingebauten Vorrang der neuen Information beinhalten, sodass diese eventuell einfach zurückgewiesen wird.

(ii) Bei Annahmen haben wir oft Gründe, diese wahr zu halten. Wir haben zumindest positives Vertrauen in ihre Wahrheit (mehr als in die ihres Gegenteils). Bei Unterstellungen brauchen wir keine besonderen Gründe. Wir lassen es offen, ob sich die Unterstellung als berechtigt erweist. Und damit räumen wir die vorrangige Revision solcher Unterstellungen durch neue Information ausdrücklich ein.

Unterstellungen sollten daher auch in Formalisierungen von Schlüssen, die solche machen, *anders* behandelt werden als Annahmen (etwa durch eine eigene Kolumne von Unterstellungen, von denen eine Zeile einer Ableitung abhängt).

Auch Wahrscheinlichkeitsschlüsse sind nicht-monoton, indem die Hinzunahme einer weiteren Information die Wahrscheinlichkeit einer Konklusion *reduzieren* kann.

(2) 1. Die meisten Hauskatzen sind nicht schwarz
(d.h. die bedingte Wahrscheinlichkeit, dass eine Hauskatze nicht schwarz ist, liegt bei $x > \frac{1}{2}$).
2. Tara ist eine von Frau Ingels vielen Hauskatzen.

$\therefore$ Tara ist nicht schwarz.

Die Konklusion wird enttäuscht durch die Zusatzinformation

(3) Frau Ingel hat fast nur schwarze Hauskatzen
(d.h. die bedingte Wahrscheinlichkeit, dass eine von Frau Ingels Hauskatzen schwarz ist, liegt bei $x > 0{,}7$).

Bemerkenswert ist hier, dass die kritische Zusatzinformation in solchen Fällen *spezifischer* ist als die vormals genutzte Information. Im Beispiel ist Tara eine Hauskatze, *weil* sie eine von Frau Ingels Hauskatzen ist (zweite Prämisse in (2)). Aufgrund dieser spezifischen Information ist (3) einschlägiger als die erste Prämisse in (2).

Als Regeln gelten hier somit:

(SS3) Benutze stets die spezifischste Information.

Dabei handelt es sich insoweit um ein Schlussprinzip als (SS3) die Auswahl von Prämissen bestimmt.

(DF1) Ist $Prob(A/B) \neq Prob(A/C)$ mit $B \supset C$, so muss eine Konditionalisierung auf Basis C im Fall B revidiert werden.

Bei (DF1) handelt es sich um eine *defeater*-Regel: die Zusatzinformation B ist ein *defeater* für den Schluss auf die neue Wahrscheinlichkeit von A aufgrund von Prob(A/C) und C.

Um diesen Umständen gerecht zu werden, wird in der Erkenntnistheorie oft die Forderung nach dem ‚Gesamtdatum' erhoben: Alle Informationen eines Subjektes sind zu berücksichtigen, sodass eine spätere Enttäuschung durch spezifischere Informationen nicht auftreten kann bzw. nur in der *Erweiterung* der Informationsbasis zu einem nachfolgenden neue Gesamtdatum auftritt. Diese Forderung idealisiert massiv: das Gesamtdatum können wir nicht überschauen. Sobald wir indessen die relevanten Informationen nach irgendwelchen Hinsichten eingrenzen, droht die Widerlegung eines schon gezogenen Schlusses durch ‚eigentlich' vorhandene *defeater* Informationen. Das Phänomen der Revision muss daher Bestandteil einer psychologisch realistischen Theorie des Schließens sein.

Bei der Nicht-Monotonie müssen eine *externe* und eine *interne Dynamik* des Schließens unterschieden werden. Eine externe Dynamik liegt typischerweise im diachronen Informationsverarbeiten vor: von t_1 zu t_2 ändert sich die Datenbasis von Δ_1 zu Δ_2. Mit dem Übergang zu Δ_2 können verbunden sein (i) der Wegfall von Antecedenzien bisher gefolgerter Konsequenzen, (ii) das Auftreten von Ausnahmebedingungen, welche bisherige Folgerungen blockieren bzw. zu deren Revision zwingen, insofern die Konsistenz des Gefolgerten gewahrt werden soll.

Bei einer externen Dynamik können alle Schlüsse *zu* einem Zeitpunkt logisch zwingend sein. Zur Modellierung des Schließens *zu* einem Zeitpunkt bedarf es dann keiner besonderen nicht-deduktiven Logik. Erst wenn die Dynamik (etwa der Wissenserweiterung und Wissensrevision) *selbst* modelliert werden soll, bedarf es zusätzlicher Modellbildung und eventuell nicht-deduktiver Logik.

Davon zu trennen ist eine interne Dynamik einer Prämissenmenge/eines Wissensstandes. Hier ergibt sich die Fluktuation in der Menge des Gefolgerten nicht durch die Berücksichtigung weiterer Informationen, sondern dadurch, dass im schrittweisen Folgern spätere Folgerungen frühere Folgerungen unterminieren. Hier muss eine nicht-monotone Logik den Prozess des Folgerns bestimmen.

8. Metatheorie

Eine metalogische Metatheorie der Prinzipien des nicht-deduktiven Schließens setzt in der Regel deren Fassung in ein formales System oder zumindest den Teil eines solchen voraus. Insbesondere bedürfte es in entsprechenden semantischen Überlegungen einer passenden Modelltheorie. Beides erweist sich als schwierig, insofern die hier diskutierten Prinzipien des nicht-deduktiven Schließens dafür auf eine gemeinsame formale Grundlage gestellt werden müssten, wobei diese sowohl modallogische als auch wahrscheinlichkeitstheoretische Elemente integrieren müsste. Insoweit hier noch nicht alle Prinzipien des nicht-deduktiven Schließens gefunden wurden, muss diese Grundlage auch entsprechend erweiterungsfähig bleiben.

Ein besonderes metatheoretisches Problem, das sich im Kontext nicht-deduktiven Schließens mehrfach stellt, nämlich immer dann, wenn man von den intuitiv qualitativen zu den mathematisch numerischen Repräsentationen übergehen will, ist das der *Metrisierung*. Es bedarf hier eine Modellbildung, die wenn schon nicht mit dem Anspruch auf psychische Realität der in ihr verwendeten Rechenverfahren, so doch zumindest mit dem Anspruch einer korrekten Implementierung der qualitativen Prinzipien auftritt. Ein Beispiel könnten baysianische Netzwerke sein.

Nicht-deduktives Schließen fließt nicht neben unserem deduktiven Schließen her, oder umgekehrt, sondern unsere Herleitungen von Konklusionen enthalten sie Schlussregeln gemischt, die gerade zur Anwendung kommen müssen. Entsprechend muss es ein allgemeines Format von Herleitungen geben. Es bietet sich an, dies im Anschluss an (deduktive) Modelle des Natürlichen Schließens zu entwickeln. Komponenten bilden:

- das deduktive Schließens mindestens einer Prädikatenlogik 1ter Stufe, das allerdings nicht aus Inkonsistentem Beliebiges folgert (also parakonsistent ist)
- das induktive und statistische Schließen mit den hier erläuterten Prinzipien (inklusive solcher Evidenzprinzipien, die auch als Schlussprinzipien dienen können)
- nicht-monotones Schließen (mit entsprechenden *default* Unterstellungen)
- evtl. weitere Komponenten wie mengentheoretisches Schließen.