Vorlesungen über

Philosophische Logik

André Fuhrmann

λογος

Vorlesungen über
Philosophische Logik

Vorlesungen über

Philosophische Logik

André Fuhrmann

Logos Verlag Berlin

λογος

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Die Open-Access-Publikation dieses Buches wurde durch den Open-Access- Publikationsfonds der Goethe-Universität Frankfurt am Main unterstützt.

ISBN 978-3-8325-5654-9

Hergestellt als Logos-Ökobuch.
https://www.logos-verlag.de/oekobuch

Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10, 12681 Berlin

Tel.: +49 (0)30 / 42 85 10 90
Fax: +49 (0)30 / 42 85 10 92

https://www.logos-verlag.de

Für Susanne und Felix & Gesine

VORWORT

Dieses Buch ist aus Vorlesungen hervorgegangen, die ich seit 2012 in jedem Sommersemester an der Goethe-Universität Frankfurt am Main gehalten habe. Diese setzten die jeweils im Winter gelesene “Einführung in die Logik” fort und waren gedacht als Angebot an besonders interessierte Studenten. Die Vorlesungen, unter dem Titel “Philosophische Logik” angekündigt, wurden mit der Zeit unter dem Namen “Sommerlogik” bekannt – im Gegensatz zur “Winterlogik” der kälteren Jahreszeit. Die wechselnden Themen der Sommerlogik entsprechen etwa den Kapiteln des Buches.

Ich danke meinen Frankfurter Studenten und Mitarbeitern, namentlich Ali Esmi, Matthias Hoch, Pascal Hohmann, Dominik Kauß, Axel Raue, und besonders Melvin Keilbar, der den gesamten Text durchgesehen und für viele Verbesserungen gesorgt hat. Ein besonderer Dank geht an Max Urchs für seine vielen Anmerkungen und Anregungen zu allen Kapiteln des Buches.

Der Universitätsbibliothek Johann Christian Senckenberg der Goethe-Universität sei für die Förderung der *Open Access*-Publikation gedankt.

Die Auswahl der Themen des Buches kann ich nicht begründen, sondern nur erklären: Es sind die Themen, zu denen ich selbst im Laufe der Jahre etwas beitragen konnte. Das Buch hätte zweifellos ein vollständigeres Bild philosophisch interessanter Logik abgeben können, wenn es einige weitere Themen eingeschlossen hätte. Allein, ...

> So eine Arbeit wird eigentlich nie fertig, man muß sie für fertig erklären, wenn man nach Zeit und Umständen das möglichste getan hat.
>
> Goethe, *Italienische Reise* (Caserta, den 16. 3. 1787)

Frankfurt am Main, den 16. 3. 2023
A.F.

INHALT

I.

EINLEITUNG

1. Philosophische Logik

Der Begriff *Logik* ist eigentlich eine Ellipse: "Logik von was?" ist die Frage, die immer sogleich beantwortet werden will. Im Falle der Aussagenlogik lautet die Antwort: Das ist die Logik der darin herausgestellten Junktoren, typischerweise der Negation, Konjunktion, Disjunktion und Implikation. In der Prädikatenlogik treten der All- und Existenz-Quantor hinzu, was erfordert, daß die Prädikationsstruktur von Aussagen freigelegt wird. Die genannten Junktoren und Quantoren sind Beispiele einer Ausdruckssorte, die manchmal "logische Partikel" (auch "logische Konstanten") genannt werden. Auf die Gefahr hin, sich im Kreise zu bewegen, könnte man sagen: Eine Logik ist eine logische Theorie bestimmter logischer Partikel. Das lädt sofort zwei weitere Fragen ein: Was ist eine logische Theorie (im Unterschied zu anderen Theorien)?, und: Was ist ein logischer Partikel (im Unterschied zu anderen Ausdrücken)?

Auf die erste Frage gibt dieses Buch eine Teilantwort, indem es Beispiele logischer Theorien vorführt. Auf die zweite Frage antworten wir hier ebenso teilweise, indem wir bestimmte Ausdrücke zum Gegenstand logischer Theorien machen. Die grundsätzliche Frage nach der "Natur" logischer Theorien und Partikel – falls sie denn eine haben sollten – ist Gegenstand der Philosophie der Logik, die wir nur gelegentlich in diesem Buch streifen werden. [1]

Nun könnte man versucht sein – ausgehend von der Frage "Logik von was?" –, den charakteristischen Gegenstand einer Philosophischen Logik

[1] Beide Fragen werden z.B. in dem Buch [294] von Alexandra Zinke eingehend behandelt. Philosophische Logik und Philosophie der Logik sind manchmal nur schwierig zu trennen. Besondes deutlich wird das im Kapitel IV über Konditionale.

so zu bestimmen: Es muß sich um etwas handeln, das in der philosophischen Theorienbildung eine wichtige Rolle spielt. Die Modallogik unter den typischen Interpretationen einer Logik der Notwendigkeit, des Glaubens, des Wissens, oder des Gebotenseins behandelt sicher zentrale theoretische Termini der Metaphysik, der Erkenntnistheorie und der Ethik. Und die Logik kontrafaktischer Konditionale trägt ebenso sicher bei zu einer philosophischen Theorie kausaler Abhängigkeit. Vermutlich ist in solchen Überlegungen der Ursprung des Begriffs 'Philosophische Logik' zu suchen. Aber eigentlich paßte der Begriff nie so recht. Denn zu den typischen Interpretationen der Modallogik gehörte beispielsweise immer auch schon die temporale. Zeitordnungen sind jedoch nicht nur Gegenstand philosophischer Theorien, sondern auch physikalischer. Das Gleiche gilt für den Begriff der Kausalität. Die Modal- und Konditionallogik sind also bestenfalls *auch* philosophische Logiken – aber nicht nur.

Vom vermutlichen Ursprung des Begriffs einer Philosophischen Logik haben wir uns jedenfalls weit entfernt. Wenn wir uns das wachsende Corpus Philosophischer Logiken ansehen, wie es beispielsweise in den zwei Auflagen des *Handbook of Philosophical Logic* [93] vorgelegt wird,[2] dann fällt auf, daß die einzelnen Bände nur durch eine *Ähnlichkkeit* der Themen oder Methoden zusammengehalten werden. Nun ist Ähnlichkeit keine transitive und also keine Äquivalenzrelation, so daß nicht sinnvoll gefragt werden kann, in welcher Hinsicht alle Philosophischen Logiken *gleich* sind und sich z.B. von der einfachen Aussagen- oder Prädikatenlogik unterscheiden. Es gibt keinen nicht-trivialen Aspekt, den diese Theorien wesentlich gemein haben. Im Übergang von der ersten zur zweiten Auflage des *Handbook* fällt weiter auf, daß sich der Anwendungsschwerpunkt zunehmend von der Philosophie in die Informatik verlagert hat.

Es ist nicht sinnvoll, nun legislativ tätig werden zu wollen, indem man den Begriff einer Philosophischen Logik normativ zurückschneidet. Der Begriff ist offen, und das ist gut so. Wir können unter den so benannten Theorien jedoch solche auswählen, die für die Belange der Philosophie von besonderem Interesse sind. Eine solche Auswahl bietet dieses Buch an.

Wir beginnen mit der *Zeitlogik* in Kapitel II. Das ist eine generische Bezeichnung für eine Familie von Modallogiken unter einer bestimmten, nämlich temporalen Interpretation. Aus rein formaler Perspektive ist diese Interpretation eine unwesentliche Heuristik. Aber wie jede gute Heuristik hilft sie, sich zu orientieren. Sie führt auch beispielhaft vor Augen, daß Philosophische Logiken etwas modellieren wollen, was außer ihnen liegt. In diesem Fall ist es das Phänomen der vergehenden Zeit. Zeitlogik ist zumin-

---

[2] Vier Bände in der ersten Auflage 1983–89, 18 Bände bis 2018 in der zweiten Auflage.

dest wesentlicher Teil einer Philosophie der Zeit.

Wir würden einen wichtigen Aspekt der logischen Theorie der Zeit verpassen, wenn wir nicht bemerken würden, wie allgemein diese Theorie tatsächlich ist. Wenn wir die theoretischen Termini einer Theorie durch die Methode der Ramsey-Sätze individuieren,[3] dann kann der Ramsey-Satz einer bestimmten Zeitlogik ununterscheidbar sein von dem einer Logik, die wir mit einer ganz anderen Heuristik versehen. Es handelt sich in dem Fall um *dieselbe* Logik, nur unter anderer (heuristischer) Interpretation ihrer theoretischen Termini. Diese Interpretation wird von der betrachteten Logik nicht festgelegt. Zeitlogik ist also nicht wesentlich Logik der *Zeit.* Im dritten Kapitel sehen wir daher von einer bestimmten Interpretation ab und gehen über zur allgemeinen *Modallogik.* Auch hier soll der Orientierung des Lesers ein wenig geholfen werden, indem wir eine Heuristik als Leitmotiv angeben: Wir lesen den Operator $\Box$ im Sinne irgendeiner Notwendigkeit. Das schließt doxastische, epistemische und deontische Notwendigkeit mit ein.

An die Modallogik schließt sich natürlich die *Konditionallogik* (Kapitel IV) an. Viele Konditionale drücken so etwas wie eine bedingte Notwendigkeit aus, so daß die Logik solcher Konditionale als eine Verallgemeinerung der Modallogik aufgefaßt werden kann. Es gibt jedoch viele Arten von Wenn-dann-Aussagen, von denen nur einige modal sind. Die Frage ihrer richtigen Einteilung und semantischen Analyse ist in hohem Grade umstritten. Das Kapitel über Konditionale versucht, den Leser durch dieses Dickicht von Theorien zu führen und einen Sinn dafür zu entwickeln, wie solche Theorien zu beurteilen sind.

In den folgenden drei Kapiteln, V–VII, geht es gewissermaßen weiter um Konditionale. Im Kapitel V über *parakonsistente Logik* betrachten wir Konditionale, die Folgerungsverhältnisse ausdrücken und versuchen diese unter die Bedingung zu stellen, daß aus widersprüchlichen Annahmen nicht Beliebiges folgt. In Kapitel VI über Relevanzlogik soll die Implikation $\rightarrow$ eine echte Beziehung zwischen Antezedens und Konsequens ausdrücken, so daß die Wahrheit von $A \rightarrow B$ nur beurteilt werden kann, wenn *beide*, A und B, aufeinander bezogen werden ("relevant" für einander sind). Die Modelltheorie der Relevanzlogik verallgemeinert die Methoden der Modal- und Konditionallogik und bietet eine sehr allgemeine Theorie konditionaler Konstruktionen an.

In Kapitel VII über *anfechtbares Schließen* geht es um nicht-deduktive Folgerungsverhältnisse. Damit sind Folgerungen gemeint, bei denen die Konklusion nicht logisch zwingend (deduktiv) aus den Prämissen folgt. Auf

[3] Siehe hierzu Lewis [180].

den ersten Blick scheint das nicht gut unter die Überschrift "Philosophische *Logik*" zu passen. Aber wir betrachten dieses Folgern (das wir generisch mit $\mid\!\sim$ bezeichnen) hier genauso wie das deduktive Folgern unter einer normativen Perspektive. D.h. wir fragen nach formalen Bedingungen, die $\mid\!\sim$ erfüllen muß, um "gerechtfertigt" zu sein. Rechtfertigung hat hier einen sehr spezifischen Sinn: Eine Relation $\mid\!\sim$ ist dann gerechtfertigt, wenn sie als vernünftiger Übergang von Prämissen zu einer Konklusion dargestellt werden kann. Im Zentrum der Theorie stehen daher Repräsentationsresultate der Art: $X \mid\!\sim A$, wenn A aus X deduktiv folgt unter kontrolliertem ("vernünftigem") Rückgriff auf bestimmte Hintergrundannahmen oder -regeln. In der hier gewählten Darstellung (die viel Makinsons [200] verdankt)ist die Theorie des anfechtbaren Schließens eine sehr allgemeine Theorie dessen, was in der Philosophie ampliatives (erweiterndes) und in der Informatik nichtmonotones Schließen genannt wird.

Wenn Kapitel VII den traditionellen Begriff einer Logik vielleicht überraschend erweitert, dann gilt das umso mehr für die *Logik des Überzeugungswandels* (Belief Revision), die Gegenstand von Kapitel VIII ist. Die Standardtheorie ist hier die AGM-Theorie der Kontraktionen und Revisionen von Überzeugungszuständen.[4] Vor die Aufgabe gestellt, unsere Überzeugungen zu ändern, müssen wir typischerweise zwischen einer Reihe von Kandidaten für unseren künftigen Überzeugungszustand wählen. Dazu müssen unsere Überzeugungszustände mit einer Struktur ausgestattet sein, die solche Wahlen ermöglichen. Wie im Falle anfechtbaren Schließens dreht sich die AGM-Theorie um bestimmte Repräsentationsresultate: Wenn die postulierte Struktur von dieser oder jener Art ist, dann haben die elementaren Änderungsoperationen der Kontraktion und Revision diese oder jene Eigenschaften. Und umgekehrt: Gegeben bestimmte formale Eigenschaften der Operationen, so dürfen wir darauf schließen, daß diese durch Strukturen der einen oder anderen Art hervorgebracht wurden. Die Theorie hat beeindruckende Anwendungen, u.a. in der Erkenntnistheorie, der Rechtslogik, der Semantik von Konditionalsätzen und der Theorie anfechtbaren Schließens.

Soweit eine kurze Beschreibung dessen, was der Leser vorfinden wird. Alle Kapitel sind weitgehend in sich abgeschlossen. Der geneigte Leser möge den unvermeidlichen Preis dafür verzeihen: daß sich nämlich zuweilen die eine oder andere Erklärung wiederholt. Querverweisen kann man überall mit Gewinn folgen, muß es aber nicht, um die Ausführungen, von denen

[4] "AGM" für: Carlos Alchourrón, Peter Gärdenfors und David Makinson.

aus verwiesen wird, zu verstehen. Alle Kapitel beginnen behutsam. Im Fortgang nimmt der technische Aufwand dann jeweils zu.

2. Grundlagen

Das Buch setzt elementare Kenntnisse der Aussagen- und Prädikatenlogik voraus. Prädikatenlogik spielt in diesem Buch allerdings nur insofern eine Rolle, als sie bequeme und eindeutige Abkürzungen zur Verfügung stellt, um gewisse Bedingungen auszudrücken. Obwohl Prädikatenlogik also nicht Gegenstand dieses Buches ist, so sollte der Leser mit ihr soweit vertraut sein, daß er die formale Sprache der Prädikatenlogik als zuweilen hilfreiche Ergänzung der Umgangssprache aufnehmen kann.

Zu den vorausgesetzten Kenntnissen im Bereich der Aussagenlogik gehören: die Beschreibung formaler Sprachen, das Ableiten von Formeln aus Annahmen oder Axiomen mittels Regeln, die Interpretation von Aussagen durch systematische (rekursive) Zuordnung von Wahrheitswerten, der Begriff einer gültigen Folgerung. Diese setzen wiederum voraus, daß der Leser mit elementaren Begriffen der Mengenlehre, einschließlich Relationen und Funktionen vertraut ist. Im folgenden seien einige wichtige Begriffe stichwortartig rekapituliert und unsere Notation erklärt.

2.1 Mengen, Folgen, Relationen, Funktionen. Objekte können in anderen Objekten vorkommen. Ein Art den Vorkommens ist die *Elementbeziehung*. Wenn x in dieser Weise in y vorkommt, dann schreiben wir $x \in y$ (bzw. $x \notin y$, wenn x kein Element von y ist). Wenn $x \in y$, dann nennen wir y eine *Menge*. Meist verwenden wir Großbuchstaben für Mengen und Kleinbuchstaben für deren Elemente (obwohl diese selbst auch Mengen sein können). Mengen werden beschrieben durch Angabe ihrer Elemente, wie in

$$\{0, 1, 2\},$$

oder durch Angabe notwendiger und hinreichender Bedingungen, wie in

$$\{x : x \text{ ist eine der ersten drei natürlich Zahlen}\}.$$

Eine Bedingung wie $x \neq x$ kann nicht erfüllt werden. Die entsprechende Menge ist daher leer; Notation $\{\,\}$ oder $\emptyset$. Eine Bedingung wie 'x ist Nachfolger einer Zahl' wird von unendlich vielen Objekten erfüllt. Die ensprechende Menge können wir mit $\{1, 2, 3, \ldots\}$ andeuten. Für bestimmte Zahlenmengen reservieren wir die folgenden Buchstaben:

$\mathbf{N}$ die natürlichen Zahlen;
$\mathbf{N}^+$ die positiven natürlichen Zahlen (> 0);

Q die rationalen Zahlen;
R die reellen Zahlen.

Mengen können in Beziehungen zueinander stehen. Insbesondere kann X *Teilmenge* von Y sein:

Teilmenge: $X \subseteq Y$ gdw jedes Element von X ist in Y, d.h. $\forall x\ (x \in X \Rightarrow x \in Y)$.

Bei Bedarf drehen wir das Zeichen $\subseteq$ um: $X \supseteq Y$ (X ist Obermenge von Y). Den Unterstrich lassen wir weg, d.h. wir schreiben $X \subset Y$, wenn X eine *echte* Teilmenge von Y ist, d.h. wenn Y Elemente enthält, die nicht in X sind. Die Menge aller Teilmengen einer Menge X ist die *Potenzmenge* von X:

Potenzmenge: $\wp(X) = \{Y : Y \subseteq X\}$.

Mengen sind unter einer Reihe einfacher Operationen abgeschlossen.

Vereinigung: $X \cup Y = \{x : x \in X \text{ oder } x \in Y\}$.
Schnitt: $X \cap Y = \{x : x \in X \text{ und } x \in Y\}$.
Subtraktion: $X \setminus Y = \{x : x \in X \text{ und } x \notin Y\}$.

Wenn $X \subseteq Y$, dann ist offenbar $X = X \cap Y$ und $X \cup Y = Y$. Die Menge X ist dieselbe Menge wie die Menge Y, wenn $X \subseteq Y$ und $Y \subseteq X$. Es folgt, daß $\{x\} = \{x, x\}$ (Idempotenz) und daß $\{x, y\} = \{y, x\}$ (Kommutativität).

Aus den Definitionen folgt unmittelbar, daß Vereinigung und Schnitt assoziative Operationen sind, d.h. $(X_0 \cup X_1) \cup X_2 = X_0 \cup (X_1 \cup X_2)$; ebenso für $\cap$. Aus diesem Grunde können wir die Klammern fortlassen und einfach $X_0 \cup X_1 \cup X_2$ schreiben. Beliebige viele Mengen können zu einer Menge vereinigt werden. Zum Beispiel können wir alle Mengen vereinigen, die eine bestimmte Bedingung ϕ erfüllen. Das Resultat würden wir so notieren: $\bigcup\{X \colon \phi(X)\}$. Häufiger werden wir die zu vereinigenden Mengen mit Indizes aus einer Indexmenge I versehen. Dann ist $\{X_i \colon i \in I\}$ die Menge der so indizierten Mengen und deren Vereinigung notieren wir mit $\bigcup\{X_i : i \in I\}$ oder auch mit $\bigcup_{i \in I}\{A_i\}$. Meist wählen wir **N** als Indexmenge, und wenn immer dies eindeutig der Fall ist, kürzen wir weiter ab zu $\bigcup A_n$. Ebenso für Schnitte $\bigcap$ von Mengen indizierter Mengen.

Wenn es uns darauf ankommt, daß Objekte in einer Ansammlung mehrfach oder in bestimmter Reihenfolge vorkommen, dann stellen wir eine solche Ansammlung nicht als Menge, sondern als *Folge* (mit gespitzten Klammern) dar. Beispielsweise kommt x in der Folge $\langle x, y, x\rangle$ zweimal vor und y geht dem zweiten Vorkommen von x voraus. Die Folgen $\langle y, x\rangle$ oder $\langle x, x, y\rangle$ würden die Anzahl und Anordnung der Vorkommen anders darstellen und sind daher verschiedene Folgen. Die Elemente einer Folge unterscheiden wir am besten, indem wir sie mit natürlichen Zahlen in ihrer natürlichen Reihenfolge indizieren. Wenn wir eine Folge $\langle x_1, \ldots, x_n \rangle$ auf diese Weise spezifizieren, dann können wir daraufhin z.B. etwas über beliebige, unmittelbar aufeinander folgende Elemente sagen, indem wir auf diese mit x_i und x_{i+1} Bezug nehmen ($1 \leq i \leq n$). Allgemein heißen Folgen mit n Elementen *n-Tupel*; im besonderen nennen wir zwei-elementige Folgen *Paare*, drei-elementige *Tripel*, usw. (*Quadrupel*, *Quintupel*, ...). Das k-te Element eines Tupels nennen wir dessen k-te *Koordinate*; im Fall eines Paares, die linke bzw. rechte Koordinate. Das *kartesische Produkt* zweier Mengen X und Y ist die Menge aller Paare $\langle x, y \rangle$ so, daß $x \in X$ und $y \in Y$. Die Verallgemeinerung liegt auf der Hand:

(Kartesisches) Produkt: $X_1 \times \cdots \times X_n = \{\langle x_1, \ldots, x_n \rangle\colon x_1 \in X_1 \;\&\; ... \;\&\; x_n \in X_n\}$.

Wenn für alle $i, k \in \{1, \ldots, n\}$ gilt, daß $X_i = X_k = X$, dann sprechen wir vom n-fachen Produkt der Menge X mit sich selbst und bezeichnen diese mit X^n. Im Grenzfall $n = 1$ sei X^1 einfach die Menge X. Statt der spitzen Klammern verwenden wir manchmal auch runde Klammern, um Folgen anzuzeigen; also beispielsweise (x, y) statt $\langle x, y \rangle$ für das Paar bestehend aus x und y in dieser Reihenfolge.

Eine zweistellige *Relation* ist eine Menge von Paaren, eine dreistellige eine Menge von Tripeln, usw. Insbesondere für zweistellige Relationen gibt es eine Reihe auffälliger und oft wiederkehrender Abschlußbedingungen. Z.B. könnte für eine Relation $R \subseteq X \times Y$ gelten, daß wenn $\langle x, y \rangle$ und $\langle y, z \rangle$ in R sind, dann ist auch $\langle x, z \rangle$ in R, für beliebige $x, y \in X$ und $y, z \in Y$. Statt $\langle x, y \rangle \in R$ schreiben wir meist kürzer Rxy (oder auch xRy) so, daß die Bedingung auch so beschrieben werden kann:

Transitivität: Wenn Rxy und Ryz, dann Rxz.

Wir nennen nun eine Auswahl weiterer solcher Bedingungen, die in diesem Buch an der einen oder anderen Stelle eine Rolle spielen werden. (Soweit nicht explizit anders gebunden, sind alle Variablen im allquantifizierten Sinne zu verstehen.)

Reflexivität:	Rxx.
Antisymmetrie:	Wenn Rxy und Ryx, dann $x = y$.
Totalität:	Rxy oder Ryx.
Symmetrie:	Wenn Rxy, dann Ryx.
Konnexität:	Rxy oder $x = y$ oder Ryx.

Dichte: Wenn Rxy, dann $\exists z$: Rxz und Rzy.
Euklidizität: Wenn Rxy und Rxz, dann Ryz.

Eine Relation $R \subseteq X^2$ ist eine *Quasiordnung* auf X, wenn R reflexiv und transitiv ist. Ist R darüberhinaus auch antisymmetrisch, dann ist R eine *Halbordnung* auf X. Ist R auch total, dann ist X *vollständig geordnet* unter R. In allen diesen Fällen schreiben wir oft $x \leq y$ statt Rxy. Eine Quasiordnung ist eine *Äquivalenzrelation*, wenn sie symmetrisch ist. In diesem Fall schreiben wir gern $x \equiv y$ für Rxy.

Wenn R dagegen irreflexiv ist, d.h. Rxx für *kein* Element der Fall ist, dann wählen wir eine Schreibweise wie $x < y$ oder $x \prec y$ für Rxy. Eine *strenge Halbordnung* ist irreflexiv, transitiv und antisymmetrisch; eine *strenge Vollordnung* ist darüberhinaus konnex.

Es sei $R \subseteq X^2$ und a ein Element a in X.

a ist *minimal* (unter R) in X: $\forall x \in X\colon Rxa \Rightarrow x = a$.
a ist *maximal* (unter R) in X: $\forall x \in X\colon Rax \Rightarrow a = x$.
a ist *das kleinste* Element (unter R) in X: $\forall x \in X\colon Rax$.
a ist *das größte* Element (unter R) in X: $\forall x \in X\colon Rxa$.

Es sei $R \subseteq X^2$ eine Halbordnung. Wenn a das kleinste Element in X ist, dann ist a auch minimal, jedoch im allgemeinen nicht umgekehrt. Ist R total, dann bildet X eine Kette unter R. Wenn die Kette nicht unendlich absteigt (aufsteigt), dann gibt es ein minimales (maximales) Element und dieses ist zugleich das kleinste (größte). (Am besten veranschaulicht man sich die Verhältnisse in Hasse-Diagrammen.)

Gegeben eine Relation $R \subseteq X \times Y$, können wir für jedes $x \in X$ den *Abschluß* von x unter R bilden: $R(x) = \{y : Rxy\}$. Wenn R eine Äquivalenzrelation ist, dann ist $R(x)$ die *Äquivalenzklasse* von x unter R; statt $R(x)$ schreiben wir dann meist $[x]_R$. Aufgrund der Eigenschaften von R gilt: $[x]_R = [y]_R$ gdw Rxy ("$x \equiv y$"), und $[x]_R \neq [y]_R$ gdw $[x]_R \cap [y]_R = \emptyset$.

Unter den Relationen $R \subseteq X \times Y$ gibt es solche, die jedem $x \in X$ höchstens ein $y \in Y$ zuordnen. Solche Relationen nennen wir *Funktionen*:

Funktionalität: Wenn Rxy und Rxz, dann $y = z$.

In diesem Fall schreiben wir $f : X \longrightarrow Y$ ("f bildet X nach Y ab") statt $R \subseteq X \times Y$ und $f(x) = y$ statt Rxy; X ist der *Argument-* oder *Definitionsbereich*, Y der *Wertebereich* von f. In $f(x) = y$ nennen wir entsprechend x das *Argument* und y den *Wert* (von x unter f). Man beachte, daß das Argument einer Funktion auch ein n-Tupel $\langle x_1, \ldots, x_n\rangle$ sein kann, in welchem Fall wir von einer *n-stelligen Funktion* sprechen und statt $f(\langle x_1, \ldots, x_n\rangle)$ vereinfacht $f(x_1, \ldots, x_n)$ schreiben. Die Bedingung der Funktionalität schließt erstens nicht aus, daß es nicht für alle Argumente $x \in X$ einen Wert in Y gibt; sie schließt zweitens nicht aus, daß f distinkte Werte in X auf denselben Wert in Y abbildet; und sie läßt drittens offen, ob jedes Element in Y sich als Wert eines Argumentes aus X unter f darstellen läßt. Im ersten Fall ist die Funktion nur *partiell*, im zweiten Fall nicht *injektiv* (umkehrbar), im dritten Fall nicht *surjektiv* (den Wertebereich ausschöpfend). Partielle Funktionen spielen in diesem Buch keine Rolle; sie können überdies in aller Regel gut durch eine Festsetzung ausgeschlossen werden. Im folgenden sei nur von Funktionen die Rede, die ihren Argumentbereich vollständig (*total*) in ihren Wertebereich abbilden. Die anderen zwei Fälle werden durch diese Bedingungen ausgeschlossen:

Injektivität	Wenn $f(x) = z$ und $f(y) = z$, dann $x = y$.
Surjektivität	$\forall y \in Y,\ \exists x \in X\colon\ y = f(x)$.

Eine (totale) Funktion $f : X \longrightarrow Y$ ist *bijektiv*, wenn f injektiv und surjektiv ist. In diesem Fall sprechen wir auch von einer eineindeutigen Abbildung oder einen Bijektion zwischen X und Y. Gibt es zwischen zwei Mengen eine Bijektion, dann sind diese *gleich mächtig*, d.h. sie enthalten die gleiche Anzahl von Elementen. Wenn X eine Menge ist, dann ist eine Bijektion zwischen X und einem (anfänglichen) Teilsegment der natürlichen Zahlen eine *Abzählung* von X.

2.2 Aussagenlogische Sprachen. Die syntaktische Beschreibung einer Sprache besteht aus der Angabe eines Alphabets sowie der Beschreibung von Regeln mit deren Hilfe aus dem Alphabet wohlgeformte Ausdrücke erzeugt werden können. Im Falle aussagenlogischer Sprachen ist das besonders einfach. Das Alphabet besteht aus einer abzählbaren Menge ATM atomarer Formeln, hier kurz *Atome* genannt, sowie einer Menge OPR von *Operatoren* (auch *Junktoren* oder “logische Konstanten” genannt) unter Angabe ihrer Stelligkeit. Da wir hier nur maximal zweistellige Junktoren betrachten, so reicht eine Stelligkeitsfunktion $f : \mathrm{OPR} \longrightarrow \{0, 1, 2\}$. Die Menge FML der wohlgeformten Ausdrücke (*Formeln*) wird durch diese Regeln definiert:

F1. $\mathrm{ATM} \subseteq \mathrm{FML}$.
F2. Wenn $x \in \mathrm{OPR}$ mit $f(x) = 0$, dann $x \in \mathrm{FML}$.
F3. Wenn $x \in \mathrm{OPR}$ mit $f(x) = 1$ und $y \in \mathrm{FML}$, dann $xy \in \mathrm{FML}$.
F4. Wenn $x \in \mathrm{OPR}$ mit $f(x) = 2$ und $y, z \in \mathrm{FML}$, dann $xyz \in \mathrm{FML}$.
F5. FML ist die kleinste Menge, die F1–4 erfüllt.

Aus der Definition folgt, daß Formeln beliebig lang sein können aber endlich sein müssen. Ob ein Ausdruck eine Formel ist, ist entscheidbar. Die Definition ist *induktiv* (oder *rekursiv*, wie man auch sagt) und erlaubt daher induktive Beweise über den Aufbau einer Formel. D.h., um zu zeigen, daß jede Formel eine bestimmte Eigenschaft hat, genügt es zu zeigen, daß 1. jedes Atom die Eigenschaft hat, und daß 2. alle Junktoren Formeln so zusammensetzen, daß die Eigenschaft von diesen an das Resultat der Zusammensetzung weitergegeben wird. Ferner sind Formeln Präfixkonstruktionen (d.h. in sogenannter polnischer Notation). Sei bespielsweise x in F4 die Konjunktion $\wedge$, dann ist nach dieser Regel $\wedge yz$ eine Formel (falls y und z Formeln sind). Die Präferenz für Präfixkonstruktionen in der offiziellen Definition von FML vereinfacht einiges, macht aber Formeln nicht besonders lesefreundlich. Deshalb werden wir uns hier auf Formeln der jeweiligen Objektsprache durch Infixausdrücke aus unserer Metasprache (angereichertes Deutsch) beziehen: also "$x \wedge y$" für "$\wedge xy$". Wir müssen dann nur durch das Setzen von Klammern für eindeutige Bezugnahme sorgen. Als (metasprachliche) Variablen verwenden wir für Atome $P, Q, R, ...$, für beliebige Formeln $A, B, C, ...$, zuweilen vermehrt und unterschieden durch Indizes aus $\mathbf{N}$. Rein syntaktisch gesehen, können die Atome auch als Operatoren mit Stelligkeit 0 aufgefaßt werden. Aber das tun wir hier nicht, da wir die Interpretation der Atome variieren wollen, während die der Operatoren in allen Modellen konstant bleiben soll. Zu den Operatoren mit Stelligkeit 0 gehören daher typischerweise nur $\top$ (*verum*) und $\bot$ (*falsum*). Mengen von Formeln bezeichnen wir meist mit $X, Y, Z, ...$; aber diese Großbuchstaben können auch – wie schon zuvor – andere Objekte oder Mengen bezeichnen. Der Kontext wird hier immer genügend Klarheit schaffen.

Mit Ausnahme der Systeme in Kap. V und VI basieren alle Überlegungen in diesem Buch auf der klassischen Aussagenlogik. Diese ist eine Theorie der Wahrheitsfunktionen in zwei Werten. Wir nehmen daher an, daß unter den Junktoren der betrachteten aussagenlogischen Sprachen sich solche befinden, die wir als eine *funktional vollständige* Menge von Wahrheitsfunktionen interpretieren können. Funktional vollständig ist eine solche Menge, wenn ihre Elemente genügen, alle Wahrheitsfunktionen zu definieren. Welche Menge das ist, ist im Prinzip gleichgültig. Wir wählen hier meist $\{\neg, \wedge\}$.

2.3 Axiomatische Systeme. Sehr allgemein betrachtet, ist ein axiomatisches System eine Struktur (S, A, R) mit einer Trägermenge S, $A \subseteq S$, und R einer Menge von Paaren $\langle X, y\rangle$ mit $X, \{y\} \subseteq S$. Der Abschluß von A unter den Elementen in R erzeugt die Menge der Theoreme des Systems. Für die Zwecke dieses Buches deuten wir die Elemente eines axiomatischen Systems spezifischer: S ist hier die Formelmenge einer aussagenlogischen *Sprache*, A ist eine entscheidbare Teilmenge von S (die *Axiome* des Systems), und R ist eine entscheidbare Menge von *Regeln*, welche *de facto* nie mehr als zwei Prämissen benötigen. Regeln sind nun also Paare $\langle X, A\rangle$ mit

$X, \{A\} \subseteq$ FML, die wir so

$$X \Rightarrow A \quad \text{oder so} \quad \frac{X}{A}$$

notieren werden.[5] In der Beschreibung eines axiomatischen Systems werden wir hier durchweg *Axiomenschemata* den Vorzug geben. Damit ist folgendes gemeint. Wenn wir z.B. festlegen, daß $A \to A$ "ein Axiom" sein soll, dann ist gemeint, daß alle Formeln von dieser Gestalt Axiome sein sollen.[6] Ohne hier "von gleicher Gestalt" zu definieren, sollte klar sein, daß wir eine adäquate Definition angeben können, nach der die Gleichgestaltigkeit beliebiger Paare von Formeln eine entscheidbare Eigenschaft ist. Ebenso verfahren wir mit der Angabe von Regeln. Mit diesem Vorgehen legen wir uns darauf fest, daß Logik "formal" ist: Wenn eine Aussage logisch wahr ist, dann ist jede Aussage gleicher Form ebenfalls logisch wahr. Jedenfalls erfüllen alle in diesem Buch behandelten Systeme diese Bedingung.

Die Menge der *Theoreme* eines axiomatischen Systems **S** ist nun so bestimmt:

T1. Jede Instanz eines Axioms von **S** ist ein Theorem von **S**.

T2. Wenn alle Formeln in einer Menge X Theoreme sind und $X \Rightarrow A$ eine Regel von **S** ist, dann ist auch A ein Theorem von **S**.

T3. Die Menge der Theoreme von **S** ist die kleinste Menge, die 1 und 2 erfüllt.

Dies ist ein weiteres Beispiel einer induktiven Definition. Um nachzuweisen, daß alle Theoreme eine bestimmte Eigenschaft haben, wird es genügen, 1. die Eigenschaft in den Axiomen nachzuweisen und dann zu zeigen, daß 2. jede Regel des Systems die Eigenschaft von den Prämissen an die Konklusion weitergibt. Das ist das Muster einer Induktion über die Definition der Menge der Theoreme eines Systems.

In einem axiomatischen System kann man *Beweise (Ableitungen) aus Annahmen* führen. Ein Beweis von B aus Annahmen $A_0, ..., A_n$ ist eine (endliche) Folge $\langle C_1, \ldots, C_n \rangle$, in der

[5] Man beachte: Die *primitiven* Regeln eines axiomatischen Systems werden in diesem Buch nie mehr als zwei Prämissen haben. Abgeleitete oder zulässige Regeln (siehe gleich die Erklärung) mögen aber im Prinzip jede endliche Anzahl von Prämissen haben.

[6] Die Alternative bestünde darin, bestimmte Formeln als Axiome (bzw. Regeln) herauszugreifen sowie eine Substitutionsregel hinzuzufügen, die aus dieser Formel alle Instanzen des entsprechenden Schemas hervorbringt.

B1. $C_n = B$, und

B2. jedes C_i ($0 \leq i \leq n$) ist entweder (a) ein Axiom, oder (b) eine der Annahmen, oder (c) eine Konklusion aus vorangehenden Formeln aufgrund einer der Regeln des Systems.

Wenn es in einem System **S** einen Beweis von B aus der Menge von Annahmen $A_0, ..., A_n$ gibt, so notieren wir das mit $A_0, ..., A_n \vdash_{\mathbf{S}} B$ (und nennen einen solchen Ausdruck eine *Sequenz*). Die Relation $\vdash_{\mathbf{S}} \subseteq \wp(\text{FML}) \times \text{FML}$ ist die *Ableitbarkeitsrelation* (oder auch Beweisbarkeitsrelation) für **S**.

BEOBACHTUNG 1. *Eine Formel A ist genau dann ein Theorem (eines Systems* **S***), wenn es einen Beweis (in* **S***) von A aus der leeren Annahmenmenge gibt.*

BEWEIS. Angenommen A ist ein Theorem. Wir wollen zeigen, daß $\emptyset \vdash A$. (Induktion über die Menge der Theoreme, s. oben.) Wenn A ein Axiom ist, dann ist $\langle A \rangle$ nach B2(a) ein Beweis von A; also $\emptyset \vdash A$. Angenommen, alle Formeln in X sind ableitbar und A ist Konklusion einer Regel mit Prämissen X. Dann gibt es für jede Formel $B \in X$ einen Beweis $\langle \ldots, B \rangle$. Wenn wir nun alle diese Beweise (in beliebiger Reihenfolge) verketten, erhalten wir eine Folge, die selbst die Bedingungen B1 und B2 erfüllt. Wenn wir der Folge schließlich A als terminales Element hinzufügen, so halten wir damit die Bedingung B2(c) ein. Also gilt auch in diesem Fall $\emptyset \vdash A$.

Sei umgekehrt $\emptyset \vdash A$. Dann ist nach B2(a,c) und T1–2 jedes Element in einem Beweis von A ein Theorem, also auch A. ∎

Statt $\emptyset \vdash A$ schreiben wir ab nun einfacher $\vdash A$. Nach der Beobachtung drücken wir damit aus, daß A ein Theorem (des gerade betrachteten Systems) ist. Wenn **S** ein axiomatisches System ist, dann werden wir die Notation auch manchmal überdehnen und mit **S** die Menge der Theoreme des Systems bezeichnen.

Wenn A ein Theorem ist, dann können wir A verwenden, um weitere Theoreme abzuleiten. Damit ist gemeint: Wenn wir die Bedingung B2 um die Klausel (d) "oder C_i ist ein Theorem" erweitern, dann wird dadurch zwar die Menge der Ableitungen, nicht aber die der Theoreme größer. Denn an jeder Stelle einer Ableitung im erweiterten Sinne, können wir nach Klausel (d) eingefügte Theoreme durch ihre Ableitungen im strengen Sinne ersetzen und erhalten so eine weitere Ableitung im strengen Sinne. Eine Formel kann also genau dann im erweiterten Sinne abgeleitet werden, wenn sie im strengen Sinne abgeleitet werden kann.

Ebenso können wir mit Regeln verfahren. Wenn wir gezeigt haben, daß $X \vdash A$, dann können wir gefahrlos nach der Regel $X \Rightarrow A$ schließen. Denn wenn wir die Definition einer Ableitung um die neue Regel erweitern, dann werden wir zwar mehr Ableitungen aber nicht mehr Sequenzen erzeugen können als zuvor. Solche Regeln heißen *ableitbare Regeln.* Davon zu unterscheiden sind *zulässige* Regeln. Die Regel $X \Rightarrow A$ ist zulässig, wenn unter der Voraussetzung, daß alle Formeln in X Theoreme sind, auch A ein Theorem ist. Also

$A_0, \ldots, A_n \Rightarrow B$ ist *ableitbar*: $A_0, \ldots, A_n \vdash B$;
$A_0, \ldots, A_n \Rightarrow B$ ist *zulässig*: wenn $\vdash A_0, \ldots, \vdash A_n$, dann $\vdash B$.

Offensichtlich ist jede ableitbare Regel auch zulässig. In der klassischen Logik gilt auch die Umkehrung. In nichtklassischen Logiken ist das jedoch nicht immer der Fall. Hier mag es vorkommen, daß wir von A auf B schließen können, falls A ein Theorem ist, jedoch nicht, falls A eine bloße Annahme ist. Beispielsweise ist die Menge der Theoreme mancher Relevanzlogiken unter der Regel $A, \neg A \vee B \Rightarrow B$ abgeschlossen, ohne daß die Sequenz $A, \neg A \vee B \vdash B$ ableitbar ist. In der intuitionistischen Logik ist z.B. die Distributionsregel $\neg A \to (B \vee C) \Rightarrow (\neg A \to B) \to (\neg A \to C)$ zulässig, jedoch nicht ableitbar. Auch in bestimmten Erweiterungen der klassischen Logik ist die Unterscheidung relevant, wie wir in Kap. III.6 sehen werden.

Beweise in einem axiomatischen System stellen wir hier meist nicht einfach als ("horizontale") Folgen von Formeln dar, sondern als ("vertikale") Listen von Formeln, in denen jeder Eintrag fortlaufend nummeriert und mit einer Anmerkung (Rechtfertigung) versehen ist. Ein Beispiel veranschaulicht das. Wir beweisen $A \to A$ aus den Axiomen (A) $A \to (B \to A)$ und (B) $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$ mit der Regel Modus Ponens (MP):

(1)	$(A \to ((A \to A) \to A)) \to$ $((A \to (A \to A)) \to (A \to A))$	(B)
(2)	$A \to ((A \to A) \to A)$	(A)
(3)	$(A \to (A \to A)) \to (A \to A)$	aus (1) und (2) durch MP
(4)	$A \to (A \to A)$	(A)
(5)	$A \to A$	aus (3) und (4) durch MP.

Für Beweise aus Annahmen bedienen wir uns des sogenannten Lemmon-Stils (nach dem Lehrbuch *Beginning Logic* (1965) von E.J. Lemmon). Danach geben wir links neben der Zeilenzahl (n) die Annahmen an, auf denen die Ableitung der Formel rechts beruht. Dann ist jeder Zeile $X\,(n)\,A$ die Aussage zu entnehmen, daß in Zeile n die Formel A aus der Annahmenmenge X abgeleitet wurde, d.h. $X \vdash A$. Annahmen werden in die Darstellung durch Zeilen der Form $A\,(n)\,A$ eingeführt. Um Platz zu sparen, zeigen wir die Annahmen durch die Zahl der Zeile an, in der sie eingeführt wurde; also

$n\,(n)\,A$ statt $A\,(n)\,A$. Der folgende, sehr kurze Beweis von $A \to A$ illustriert das

1	(1)	A	Annahme
	(2)	$A \to A$	aus (1) und dem Deduktionstheorem (s.u.)

Nachdem das Deduktionstheorem im nächsten Abschnitt gezeigt wurde, überzeuge sich der Leser davon, daß der Schluß im Beweis richtig ist.

2.4 Folgerung: Operationen und Relationen. Theorien. Wir werden oft Anlaß haben die Menge der Formeln zu betrachten, die aus einer gegebenen Menge X logisch folgt. Der Abschluß einer Menge unter einer Operation ist ein sehr allgemeines Phänomen, welches durch die sogenannten *Tarski-Bedingungen* beschrieben werden kann. Uns interessiert hier eine solche Operation nur, insoweit sie für die Formelmengen einer Sprache definiert ist. Wir definieren daher Cn als eine Abbildung $\wp(\mathrm{FML}) \longrightarrow \wp(\mathrm{FML})$, welche die folgenden (Tarski-) Bedingungen erfüllt:

Inklusion:	$X \subseteq \mathrm{Cn}(X)$.
Idempotenz:	$\mathrm{Cn}(\mathrm{Cn}(X)) \subseteq \mathrm{Cn}(X)$.
Monotonie:	Wenn $X \subseteq Y$, dann $\mathrm{Cn}(X) \subseteq \mathrm{Cn}(Y)$.

Erfüllt Cn darüber hinaus die Bedingung

Endlichkeit	Wenn $A \in \mathrm{Cn}(X)$, dann gibt es eine endliche Menge $X' \subseteq X$ so, daß $A \in \mathrm{Cn}(X')$,

dann heißt Cn *endlich* (oder *kompakt*).

BEOBACHTUNG 2. *Es sei* $\vdash$ *die Ableitungsrelation eines axiomatischen Systems. Wenn* $\mathrm{Cn}_{\vdash}(X) = \{A : X \vdash A\}$, *dann ist* $\mathrm{Cn}_{\vdash}$ *eine endliche Konsequenzoperation.*

BEWEIS. Wir übersetzen die Tarski-Bedingungen nach der Definition von $\mathrm{Cn}_{\vdash}$ in Eigenschaften von $\vdash$ und prüfen dann, ob $\vdash$ diese Eigenschaften hat. Inklusion und Monotonie werden zu

Reflexivität:	$A \in X \;\Rightarrow\; X \vdash A$ und
Monotonie:	$(X \vdash A$ und $X \subseteq Y) \;\Rightarrow\; Y \vdash A$,

was beides offensichtlich auf $\vdash$ zutrifft. Idempotenz wird zu

$$\text{Schnitt:} \qquad (Y \vdash A \text{ und } \forall B \in Y\colon X \vdash B) \;\;\Rightarrow\;\; X \vdash A.$$

Angenommen wir haben einen Beweis δ von A aus Y. Die zweite Annahme erlaubt uns für jedes $B \in Y$, welches in δ vorkommt, eine Ableitung von B aus X einzufügen. Das Resultat ist eine Ableitung von A aus X. Da

Beweise endliche Folgen sind, kann in ihnen nur auf endlich viele Formeln einer Annahmenmenge X zugegriffen werden; also folgt auch Endlichkeit:

Endlichkeit — Wenn $X \vdash A$, dann gibt es eine endliche Menge $X' \subseteq X$ so, daß $X' \vdash A$. ■

Die Durchsicht des Beweises zeigt, daß wenn

$$A \in \mathrm{Cn}(X) \text{ gdw } X \vdash A,$$

dann erfüllt Cn die Tarski-Bedingungen genau dann, wenn $\vdash$ die im Beweis genannten Bedingungen erfüllt. Es folgen einige Beobachtungen über Cn, von denen wir häufiger Gebrauch machen werden und die leicht zu verifizieren sind:

$$\begin{aligned} \mathrm{Cn}(X) &= \mathrm{Cn}(\mathrm{Cn}(X)). \\ \mathrm{Cn}(X) \cup \mathrm{Cn}(Y) &\subseteq \mathrm{Cn}(X \cap Y). \\ \mathrm{Cn}(X \cup Y) &= \mathrm{Cn}(\mathrm{Cn}(X) \cup \mathrm{Cn}(Y)). \\ \mathrm{Cn}(X \cap Y) &\subseteq \mathrm{Cn}(X) \cap \mathrm{Cn}(Y). \\ X \subseteq \mathrm{Cn}(Y) \;\&\; Y \subseteq \mathrm{Cn}(X) \text{ gdw } \mathrm{Cn}(X) &= \mathrm{Cn}(Y). \end{aligned}$$

Die bisherigen Überlegungen waren unabhängig von der Präsenz bestimmter Junktoren in der Sprache. Zumindest eine Eigenschaft von Cn bzw. $\vdash$ *vis-à-vis* der Implikation $\to$ sei hier jedoch erwähnt, da wir ihr in einigen Kapiteln besondere Aufmerksamkeit schenken werden.

Deduktionstheorem 3. (Herbrand 1930) *In jeder Logik mit den Theoremen* (A) $A \to (B \to A)$ *und* (B) $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$, *sowie Modus Ponens* $(A, A \to B) \Rightarrow B$ *als einziger Ableitungsregel gilt: Wenn* $X, A \vdash B$, *dann* $X \vdash A \to B$.

Beweis. Es sei vorab daran erinnert, daß in jeder Logik mit den Theoremen (A) und (B) und Modus Ponens (MP) auch $A \to A$ ein Theorem ist (s. den Beweis oben). Ferner folgt aus $\vdash A$ per Monotonie $X \vdash A$.

Wenn $X, A \vdash B$, dann gibt es einen Beweis $\delta = \langle C_1, \ldots, C_n \rangle$ von B aus $X \cup \{A\}$. Wir zeigen durch vollständige Induktion über die Länge n von δ, daß es einen Beweis von $A \to B$ aus X gibt.

Im Basisfall $n = 1$ ist $\delta = \langle B \rangle$ und B ist entweder (a) ein Axiom oder (b) in $X \cup \{A\}$. Im Fall (a) haben wir $X \vdash B$. Aus Theorem (A), $\vdash B \to (A \to B)$, folgt per MP, $X \vdash A \to B$. Im Fall (b) ist $B \in X$ oder $B = A$. Der erste Unterfall ist wie Fall (a). Für den zweiten Unterfall haben wir (s.o.) $X \vdash A \to A$, d.h. $X \vdash A \to B$.

Wir gehen nun aus von der Induktionsannahme (IA), daß die Behauptung für alle δ bis Länge $k-1$ gilt. D.h., für alle k: Wenn $X, A \vdash C_{k-1}$, dann $X \vdash A \to C_{k-1}$. Wir nehmen ferner an, daß $X, A \vdash C_k$ (und zeigen, daß $X \vdash A \to C_k$). C_k ist entweder (a) ein Axiom, oder (b) in $X \cup \{A\}$, oder (c) Konklusion per MP aus vorhergehenden Formeln. Die Fälle (a–b) sind wie im Basisfall. Im Fall (c) gibt es aus $X \cup \{A\}$ ableitbare Formeln C_i und C_j $(i, j < k)$ so, daß $C_i = C_j \to C_k$. Nach der IA haben wir daher

$$X \vdash A \to C_j \text{ und } X \vdash A \to (C_j \to C_k).$$

Nach Theorem (B) und MP folgt daraus $X \vdash A \to C_k$, wie gewünscht. ∎

Man beachte, daß der Beweis in zwei Richtungen empfindlich ist: Er läßt sich so nicht führen, wenn entweder eines der beiden Theoreme nicht zur Verfügung steht, oder wenn neben Modus Ponens weitere Regeln zu berücksichtigen sind.

Unter einer *Theorie* verstehen wir eine Formelmenge T, die unter einer Konsequenzoperation Cn abgeschlossen ist; also $T = \mathrm{Cn}(T)$, oder relational ausgedrückt: $A \in T \Leftrightarrow T \vdash A$. In der Regel wird von Theorien die Rede sein, die unter der Konsequenzoperation einer *bestimmten* Logik $\mathbf{L}$ abgeschlossen ist, d.h. so, daß die Theorie T genau die Formeln enthält, die sich in $\mathbf{L}$ aus ihr ableiten lassen. In diesen Fällen sprechen wir von einer $\mathbf{L}$-Theorie. Für die meisten der hier betrachteten Kandidaten für $\mathbf{L}$ wird der Schnitt zweier $\mathbf{L}$-Theorien selbst eine $\mathbf{L}$-Theorie sein, d.h.

$$\mathrm{Cn}(X) \cap \mathrm{Cn}(Y) = \mathrm{Cn}(\mathrm{Cn}(X) \cap \mathrm{Cn}(Y)).$$

Im Zusammenhang mit Strukturen und Modellen wird der Begriff 'Theorie' noch in einem etwas anderen Sinne gebraucht werden.

Vom nächsten Lemma werden wir in diesem Buch häufig Gebrauch machen. Wir möchten sagen, eine Formelmenge sei konsistent, wenn aus ihr kein Widerspruch abgeleitet werden kann. Diese Bestimmung setzt voraus, daß eine Ableitbarkeitsrelation gegeben ist, und daß es in der betrachteten Sprache eine Formel gibt, die einen Widerspruch ausdrückt. Wir wollen eine solche Formel hier mit $\perp$ bezeichnen. Es sei $\vdash$ eine Ableitbarkeitsrelation. Dann können wir definieren:

X ist ($\vdash$-) *konsistent* gdw $X \nvdash \perp$.

X ist ($\vdash$-) *maximal* gdw für alle Formeln A gilt:
wenn $A \notin X$, dann $X, A \vdash \perp$.

LINDENBAUM-LEMMA 4. *Jede konsistente Formelmenge läßt sich zu einer maximal konsistenten erweitern.*

BEWEIS. Sei X konsistent (unter $\vdash$). Wir wählen eine beliebige Abzählung der Formelmenge FML einer Sprache und drücken mit $A_1, A_2, A_3, \ldots$ aus, daß A_1 die erste, A_2, die zweite usw. Formel unter dieser Abzählung ist. Nun definieren wir eine Folge von Mengen:

$$X_0 = X$$
$$X_{n+1} = \begin{cases} X_n \cup \{A_{n+1}\}, \text{ falls } X_n, A_{n+1} \nvdash \bot; \\ X_n \text{ anderenfalls} \end{cases}$$
$$X^* = X_0 \cup X_1 \cup X_2 \cup \cdots \ \text{(d.h. } X^* = \bigcup X_n\text{)}$$

X^* *ist konsistent.* Wir zeigen durch Induktion über n (≥ 0), daß jedes X_n konsistent ist. Für $n = 0$ gilt das per Annahme. Die Induktionsannahme (IA) sei nun, daß X_n konsistent ist. Wir zeigen, daß dann auch X_{n+1} konsistent ist. Nun ist X_{n+1} entweder X_n oder $X_n \cup \{A_{n+1}\}$. Im ersten Fall folgt die Konsistenz sofort nach der IA. Im zweiten Fall haben wir nach der Definition $X_n, A_{n+1} \nvdash \bot$, d.h. X_{n+1} ist konsistent.

X^* *ist maximal.* Angenommen, es gibt ein $A \notin X^*$ mit $X^*, A \nvdash \bot$. Die Formel A muß als A_{i+1} in der Abzählung vorkommen. Da $A_{i+1} \notin X^*$, so auch $A_{i+1} \notin X_{i+1}$ [$\subseteq X^*$]. Es folgt aus der Definition, daß $X_i, A_{i+1} \vdash \bot$. Aber dann auch $X^*, A_i + 1 \vdash \bot$ – im Widerpruch zu unserer Annahme. ∎

Maximal konsistente Mengen X sind Theorien, d.h.

$$A \in X \text{ gdw } X \vdash A.$$

Von links nach recht ist das trivial. Für die andere Richtung nehme man an, daß (1) $X \vdash A$ und (2) $A \notin X$. Dann folgt aus (2) aufgrund der Maximalität, daß (3) $X, A \vdash \bot$. Aus (1) und (3) folgt nach der Schnitteigenschaft von $\vdash$, daß $X \vdash \bot$ – im Widerspruch zur Konsistenz von X.

In Sprachen mit einer Negation $\neg$ dürfen wir erwarten, daß diese Einführungsregel gilt:

$\neg$ Ein $$X, A \vdash \bot \ \Rightarrow \ X \vdash \neg A.$$

Dann können wir zeigen, daß maximal konsistente Mengen *negationsvollständig* sind, d.h.

$$\text{wenn } X \nvdash A, \text{ dann } X \vdash \neg A.$$

Denn wenn $X \nvdash A$, dann $A \notin X$ und weiter $X, A \vdash \bot$, da X maximal ist. $\neg$ Ein ergibt dann $X \vdash \neg A$.

Diese und weitere Eigenschaften maximal konsistenter Mengen geben die Schlüssel zu den Vollständigkeitsresultaten in den folgenden Kapiteln ab.

2.5 Strukturen und Modelle. Die Sätze einer Sprache sagen nichts aus, solange sie nicht interpretiert sind. Einen Satz interpretieren heißt, Bedingungen anzugeben, unter denen er etwas richtig wiedergibt. Ein Satz gibt etwas richtig wieder, wenn er wahr ist in Bezug auf dieses Etwas. Das "Etwas" liegt typischerweise außerhalb der Sprache; es ist das, worauf der Satz sich in einem sehr allgemeinen Sinne bezieht. Wir nennen es hier eine *Struktur* oder auch einen (Bezugs-) *Rahmen*. Ob ein Satz Aspekte einer Struktur wahrhaftig wiedergibt, hängt davon ab, *wie* er auf die Struktur bezogen wird. Eine Art und Weise, Sätze auf eine Struktur zu beziehen, nennen wir eine *Interpretation*. Jede Interpretation der Sätze einer Sprache in einer Struktur ist ein *Modell* der Sprache (auf dieser Struktur).

Modelle sollen zwei Bedingungen erfüllen. Jede Interpretation einzelner Sätze soll zu einer Interpretation *aller* Sätze erweitert werden können. Wie das im einzelnen geschieht, ist meist gleichgültig. Wesentlich ist nur, daß wir Interpretationen als *totale* Funktionen von der Menge der Sätze einer Sprache in Strukturelemente auffassen: Eine Interpretation muß alle Sätze der Sprache erfassen – auch wenn die meisten davon uns nicht interessieren. Tarski nennt das die Bedingung der "materialen Adäquatheit" (und stellt fest, daß diese Bedingung unter bestimmten Umständen nicht konsistent erfüllt werden kann.)

Zweitens sollen Interpretationen "systematisch" sein. Damit ist zumindest gemeint, daß es eine endliche Darstellung geben soll, wie jeder Satz der unendlich vielen Sätze einer Sprache in endlicher Zeit interpretiert, d.h. mit einer Wahrheitsbedingung versehen werden kann. Es ist offensichtlich, daß natürliche Sprachen diese Bedingung erfüllen, denn anderenfalls wären sie nicht erlernbar. Im Semantik genannten Teil der Aussagenlogik wird diese Bedingung auf eine bestimmte Weise realisiert: Die Definition einer Interpretation rekapituliert die induktive Definition einer Formel. Im Fall einer aussagenlogischen Sprache mit Junktorenmenge $\{\neg, \wedge\}$ erlaubt die Definition F1–5 der Formelmenge die folgende Beobachtung:

> Jede Formel ist entweder ein Atom, oder eine Negation $\neg A$, oder eine Konjunktion $\wedge AB$. (Eindeutige Lesbarkeit.)

Aus der Beobachtung folgt, daß eine Funktion f – also insbesondere eine Interpretation – für den gesamten Bereich der Formeln definiert ist, wenn f für alle Atome und für die beiden Fälle $\neg A$ und $A \wedge B$ definiert ist (Definition durch Rekursion über den Formelaufbau). Die Beobachtung ist stabil unter Erweiterungen der Sprache um weitere Junktoren (oder der Wahl anderer Junktoren), vorausgesetzt diese fügen sich in eines der Schemata F2–4. [7]

[7] Das ist eine *hinreichende* Bedingung, die alle Junktoren in diesem Buch erfüllen werden.

Die einfachste Struktur, die wir in diesem Buch betrachten werden, ist eine *Boolesche Algebra* $\mathcal{B} = (B, -, \sqcap)$, wobei B eine Menge mit mindestens zwei Elementen ist, die unter Funktionen $- : B \longrightarrow B$ und $\sqcap : B^2 \longrightarrow B$ abgeschlossen ist. Wir schreiben $\bar{x}$ für $-x$ und 0 für $x \sqcap \bar{x}$. Ein Element $1 \in B$ sei definiert als $\bar{0}$. Die Funktionen $-$ und $\sqcap$ stehen unter diesen Bedingungen:

$$x \sqcap y = y \sqcap x$$
$$(x \sqcap y) \sqcap z = x \sqcap (y \sqcap z)$$
$$x \sqcap \bar{y} = 0 \text{ gdw } x \sqcap y = x$$

Wir wollen zunächst nur eine *bestimmte*, zwei-elementige Boolesche Algebra

$$\mathbf{2} = (\{a, b\}, -, \sqcap)$$

betrachten. Es folgt sogleich, daß $a = \bar{b}$ und $\bar{a} = b$, und daß $a = 0$ gdw $b = 1$. Wir können also a und b gleich mit 0 und 1 bezeichnen (oder umgekehrt). Ferner können wir uns davon überzeugen, daß in **2** gilt:

$$-x = 1 \text{ gdw } x = 0$$
$$x \sqcap y = 1 \text{ gdw } x = 1 = y$$

– was natürlich an vertraute Wahrheitstafeln erinnern sollte. Eine Funktion I interpretiert eine Sprache mit Atomen $P_0, P_1, \ldots$ und Junktoren $\neg$ und $\wedge$ in **2**, falls

B1. $\forall P \in \text{ATM}$: $I(P) \in \{0, 1\}$, und

B2. $I(\neg) = -$ und $I(\wedge) = \sqcap$.

Nach der oben erwähnten Beobachtung, gewährleisten B1–2 die Interpretation *aller* Formeln. Man beachte, daß B2 den Wert (die Bedeutung) der Zeichen $\neg$ und $\wedge$ für jede Interpretation festlegt, während B1 die Interpretationen nur unter eine Bedingung stellt, die auf verschiedene Weise erfüllt werden kann. Beispielsweise erlaubt B1 sowohl $I(P_0) = 1$ als auch $I'(P_0) = 0$. Beide Funktionen sind gute Interpretationen in **2**, vorausgesetzt B2 wird beachtet, d.h. $I(\neg P_0) = 0$ als auch $I'(\neg P_0) = 1$ usw. für alle weiteren Zusammensetzungen von P_0. In diesem Sinne ist die Interpretation der Zeichen $P_0, P_1, \ldots$ *variabel*, während die von $\neg$ und $\wedge$ *konstant* ist. (Man sagt auch: Die einen Zeichen seien die "Variablen" der Sprache, die anderen die "logischen Konstanten".)

Anmerkung. In den folgenden Kapiteln werden wir dem Usus folgen, die (variable) Interpretation der Atome von der (konstanten) Interpretation der Junktoren deutlich zu trennen. Eine Funktion I wird nur den

Atomen Wahrheitswerte zuweisen – mal in der einen, mal in der anderen Weise. Auf der Basis einer solchen Interpretation I der Atome wird dann eine *Erfüllungsrelation* (auch "Wahrmacherrelation" genannt) $\models$ über den Formelaufbau definiert. Diese Definition ist für alle Modelle gleich und beginnt mit der Festsetzung ($\forall P \in \mathrm{ATM}$), daß $\models P$ gdw $I(P) = 1$, bzw. mit einer Variation über diese Äquivalenz, je nach der Komplexität der betrachteten Strukturen.

Wenn I die Bedingungen B1–2 erfüllt, dann ist $\mathcal{M} = (\{0, 1\}, -, \sqcap, I)$ ein Modell einer Sprache vom gerade betrachteten Typ auf der Struktur **2**. Wir wollen einen Satz A *wahr in diesem Modell* $\mathcal{M}$ nennen (Notation: $\mathcal{M} \models A$), falls A unter der Interpretation des Modells den Wert 1 erhält. Wahrheit in einem Modell ist konkrete Wahrheit: Sie hängt ab von der Wahl einer konkreten Struktur und einer konkreten Interpretation. Logische Wahrheit abstrahiert von solchen Wahlen. Wir wollen nur *Arten* von Strukturen betrachten, und von den Zufälligkeiten einzelner Interpretationen sehen wir ab, indem wir nur die Interpretation der Junktoren festhalten und alles andere variieren. Wenn wir diese beiden Bedingungen einhalten, dann haben wir die Logik dieser Junktoren – die Menge der sie betreffenden logischen Wahrheiten – in einer (hoffentlich) genügend weiten Klasse von Strukturen angegeben.

In einem ersten Schritt abstrahieren wir daher von den Zufälligkeiten einzelner Modelle (d.h. deren Interpretationsfunktion) und definieren Wahrheit in der Struktur **2**, gleichgültig wie wir die Sätze in dieser *bestimmten* Struktur interpretieren:

1. A ist wahr in **2** (Notation: $\mathbf{2} \models A$) gdw $\forall \mathcal{M}$ (auf **2**): $\mathcal{M} \models A$.

Nun ist **2** eine Struktur mit einer *bestimmten* Grundmenge (bestehend aus zwei konkreten Elementen). Auf die Wahl einer solchen Grundmenge soll es aber nicht ankommen. Also betrachten wir als nächstes die Menge $\mathbb{B}_2$ aller *zwei-elementigen* Booleschen Algebren. Dann ist

2. A wahr in $\mathbb{B}_2$ (Notation: $\mathbb{B}_2 \models A$) gdw $\forall \mathcal{B} \in \mathbb{B}_2$: $\mathcal{B} \models A$.

Schließlich könnten wir noch einen Schritt weiter gehen und auch von der Mächtigkeit der Grundmenge absehen und Wahrheit in der Klasse $\mathbb{B}$ *aller* Boole'schen Algebren betrachten wollen:

3. A ist wahr in $\mathbb{B}$ gdw $\forall \mathcal{B} \in \mathbb{B}$: $\mathcal{B} \models A$.

In welchem dieser drei Schritte haben wir nun einen plausiblen Kandidaten für logische Wahrheit erreicht? Es stellt sich heraus, daß in diesem Fall (im Falle anderer Strukturen jedoch nicht) die Wahrheitsbegriffe in allen drei

Schritten koinzidieren so, daß der erste Abstraktionsschritt schon ausreicht. Danach ist eine Formel logisch wahr, wenn sie unter allen Interpretationen in der Booleschen 2-Elemente-Struktur wahr ist. Die Elemente – wie immer sie konkret aussehen mögen – können wir uns einfach als Repräsentanten der beiden Wahrheitswerte Wahr (1) und Falsch (0) vorstellen.

Wir haben dieses Beispiel etwas eingehender behandelt, weil es in einfacher Weise die zentralen Begriffe einer Struktur, einer Interpretation und eines Modells einführt. Die Strukturen, die in diesem Buch die Hauptrolle spielen, werden jedoch keine algebraischen, sondern Mengenstrukturen sein. Das sind Strukturen $(W, R_1, R_2, ...)$ mit einer Trägermenge W und einer Reihe von Relationen (manchmal auch Funktionen) auf W. Die Elemente in W werden zuweilen "Welten" genannt; in der Zeitlogik werden sie als Zeitpunkte interpretiert. Die neutralen Bezeichnungen *Punkte* oder *Indizes* passen in jedem Fall. Ihre Hauptaufgabe ist es, die Interpretation von Formeln zu relativieren. Eine Interpretation auf einer solchen Struktur verteilt Wahrheitswerte über die Formeln immer nur relativ zu Indizes. Wie das im Detail geschieht, wird in den einzelnen Kapiteln erklärt.

Unabhängig von der Natur der Strukturen, können wir hier einige Begriffe einführen, die in vier der folgenden Kapitel eine wichtige Rolle spielen werden. Diese Begriffe fassen bestimmte Beziehungen zwischen den Formeln einer gegebenen Sprache und Strukturen bestimmter Art zusammen.

$Th(\mathbb{S}) = \{A : \mathbb{S} \models A\}$ nennen wir die *Theorie der Strukturklasse* $\mathbb{S}$, d.h. die Menge der in $\mathbb{S}$ wahren Sätze der betrachteten Sprache.[8]

$Rm(X) = \{\mathcal{S} : \mathcal{S} \models A, \forall A \in X\}$ nennen wir die Klasse der *Strukturen für die Formelmenge* X, d.h. die Klasse aller Strukturen, welche die Sätze in X wahr machen.

X ist *richtig* bezüglich $\mathbb{S}$: $X \subseteq Th(\mathbb{X})$.
X ist *vollständig* bezüglich $\mathbb{S}$: $Th(\mathbb{S}) \subseteq X$.

X *definiert* die Strukturklasse $\mathbb{S}$: $Rm(X) = \mathbb{S}$.

Sei A ein Formelschema und (Bed) eine Bedingung auf einer Strukturklasse $\mathbb{S}$. Wir notieren mit $Bed(\mathbb{S})$ die Strukturen in $\mathbb{S}$, welche die Bedingung erfüllen. Dann *korrespondiert* A mit (Bed) gdw $Rm(A) = Bed(\mathbb{S})$ ("Das Schema definiert die Bedingung").

Die Richtigkeit und Vollständigkeit der klassischen Aussagenlogik **KL** in Bezug auf die Booleschen Algebren, die wir oben betrachtet haben, können

[8] Dieser Gebrauch des Wortes "Theorie" ist unabhängig von der oben eingeführten Rede von Theorien als unter Konsequenz abgeschlossenen Mengen. Durch die Einführung einer geeigneten Abschlußoperation kann in bestimmten Fällen eine Brücke hergestellt werden.

wir dann beispielsweise so festhalten:

$$Th(\mathbf{2}) = Th(\mathbb{B}_2) = Th(\mathbb{B}) = \mathbf{KL}.$$

2.6 Wahrscheinlichkeit. Grundwissen über den Begriff der Wahrscheinlichkeit gehört ebenso zum Handwerkszeug des Philosophen wie Mengentheorie. Da wir davon aber nur im Kapitel IV (über Konditionale) Gebrauch machen werden, werden die nötigen Grundlagen dort (im Abschnitt 7) bereitgestellt.

3. Notation

Wir benutzen gebräuchliche Abkürzungen – wie "gdw" (genau dann wenn), "LR" bzw. "RL" (von links nach rechts bzw. umgekehrt), "zz" (zu zeigen, in einem Beweis), "IA" (Induktionsannahme), "■" (Beweisende), und einige weitere –, die dem Leser keine Rätsel aufgeben sollten.

Technische Notation wird bewußt zurückhaltend eingeführt, d.h. nur dort, wo sie dem Verständnis dient und nur mit der Genauigkeit, die der jeweilige Kontext erfordert. Relativierende Bezüge, angezeigt durch Indizes, lassen wir immer dort weg, wo der Bezug im Kontext unübersehbar ist. In jedem Fall wird auf solche Vereinfachungen hingewiesen. Es folgen stichwortartige Erläuterungen, etwa in der Reihenfolge der Einführung der Notation.

Sprachen

werden mit $\mathcal{L}$ bezeichnet, manchmal zur Unterscheidung mit Sub- oder Superskript. ATM, LIT, OPR, FML sind die Mengen der atomaren Formeln, der Literale (Atome oder deren Negationen), der Operatoren (Junktoren), und der Formeln einer gegebenen Sprache.

$P, Q, R, ...$	atomare Formeln
$A, B, C, ...$	Formeln
$X, Y, Z, ...$	Mengen von Formeln
$\mathrm{tf}(A)$	Menge der Teilformeln von A

Logiken

werden immer durch **F**ettbuchstaben bezeichnet. Logiken werden hier meist axiomatisch definiert und bezeichnen dann in engerer Verwendung das jeweilige axiomatische System, in weiterer Verwendung jedoch auch die Menge der Theoreme eines solchen Systems.

$\vdash$	Theorem bzw. Ableitbarkeitsrelation (in einer Logik)
KL	klassische Aussagenlogik
PL	klassische Prädikatenlogik
IL	intuitionistische Aussagenlogik
JL	Johanssons Minimalkalkül
$\mathbf{L}$+S, $\mathbf{L}$S	Erweiterung von $\mathbf{L}$ um Schema S

Kapitel II (Zeitlogik)

In diesem, wie auch in den folgenden Kapiteln stehen kalligraphische Buchstaben $\mathcal{R}, \mathcal{S}, \mathcal{M}, ...$ für Strukturen/Rahmen bzw. Modelle; Mengen von Strukturen bzw. Modellen werden durch doppelt gestrichene Buchstaben $\mathbb{R}, \mathbb{S}, \mathbb{M}, ...$ bezeichnet. Die Relation $\models$ nimmt je nach Kontext eine besondere Bedeutung an, die letztlich immer auf der grundlegenden Bedeutung von $a \models_{\mathcal{M}} A$ fußt: der Punkt a macht im Modell $\mathcal{M}$ die Formel A wahr.

T	Menge von Zeitpunkten
$a, b, ..., s, t, ...$	
$..., x, y, ...$	Zeitpunkte
$\prec, \succ$	Relationen zwischen Zeitpunkten
P	irgendwann in der Vergangenheit
F	irgendwann in der Zukunft
H	immer in der Vergangenheit
G	immer in der Zukunft
$\otimes$	disjunkte Vereinigung zweier Strukturen
A^d	Dual der Formel A
A^s	Spiegelbild der Formel A
seit	Seit-Operator
bis	Bis-Operator
@	Jetzt-Operator
$ab \models$	wahr an a bzgl. b
$\mathbf{K}$	minimale Zeitlogik (mit F)
$\mathbf{K}_0$	minimale Zeitlogik (mit F und P)
$\mathbf{K}_t$	minimale Zeitlogik mit Spiegelung
$\mathbf{K}_t^*$	“richtige” Zeitlogik
$\Box, \Diamond$	geschichtliche Notwendigkeit/Möglichkeit

Kapitel III (Modallogik)

Hier werden einige Gegenstände aus der Zeilogik teilweise in geänderter Bezeichnung erneut eingeführt. Statt von Strukturen $(T, \prec)$ sprechen wir z.B. jetzt vorzugsweise von Rahmen (W, R), mit geänderter Bezeichnung der Grundmenge und der Zugangsrelation.

W	Menge von Punkten ("Welten")
$a, b, ..., x, y, ...$	Punkte
R	Zugangsrelation zwischen Punkten
$[\![A]\!]$	Menge aller A-Welten (Proposition)
$\Box, \Diamond$	Notwendigkeit/Möglichkeit
$\twoheadrightarrow$	strikte Implikation
$\circ$	Verträglichkeit
$Th(\mathbb{R})$	Theorie der Rahmenklasse $\mathbb{R}$
$Mod(\mathcal{R})$	Modelle auf Rahmen $\mathcal{R}$
$\mathcal{M}^{\mathbf{L}}$	kanonisches Modell für Logik **L**
$\models^{\ell}, \models^{g}$	lokale/globale Folgerung
$Rm(X)$	Rahmen für Formelmenge X
$Bed(\mathbb{K})$	Kripke-Rahmen unter Bedingung (Bed)
$\mathcal{M}^{a}$	aus Punkt a erzeugtes Teilmodell
$\otimes$	disjunkte Vereinigung zweier Rahmen
Bew	Beweisbarkeitsprädikat
$\ulcorner A \urcorner$	Gödelzahl von A
G, L, H	Fixpunktformeln
$[a]_A$	Äquivalenzklasse von a bzgl. der Teilformeln von A
$\mathcal{M}_A$	Filtrat von $\mathcal{M}$ für A
(W, f)	funktionaler Rahmen
(W, N)	Nachbarschaftsrahmen

Für Modallogiken folgen wir weitgehend der Benennungskonvention von Chellas [53] (s. dazu die Anmerkung in Kapitel 3, Abschn. 4).

K	kleinste normale Modallogik
KT	= **T**
KT4	= **S4**
KT5	= **S5**
KD	= **SDL**, Standardsystem der deontischen Logik
KW	= **GL** (Beweisbarkeitslogik)
PA	formalisierte Arithmetik (mit Induktion)
E	kleinste kongruente Modallogik

Kapitel IV (Konditionale)

$\sqsupset$	konjunktivisches "wäre"-Konditional
$>$	konjunktivisches "könnte"-Konditional
$R_{[\![A]\!]}, R_X$	Zugangsrelation (für $\sqsupset$ und $>$)

$\mathfrak{S}$	Sphärenfunktion
$\mathfrak{S}_a$	Sphärenmenge zentriert auf a
S, S', S''	Sphären
$\min_a$	minimale Sphäre (in $\mathfrak{S}_a$)
μ_a	minimaler Schnitt (in $\mathfrak{S}_a$)
$\succ$	indikativisches Konditional
p	Wahrscheinlichkeit
p_B	bedingte Wahrscheinlichkeit
$\overline{\mathrm{p}}$	Unwahrscheinlichkeit (Ungewißheit)
$\overline{\mathrm{p}}_B$	bedingte Unwahrscheinlichkeit
$\models_{\mathrm{p}}$	Wahrscheinlichkeitsfolgerung
$\triangleright$	Ableitbarkeit in Adams' System
CK	kleinste normale Konditionallogik
V	**CK** + unkontroverse Postulate
VW	**V** mit schwacher Zentrierung
VC	**V** mit starker Zentrierung (Lewis-Logik)
VCS	**VC** mit Ausgeschlossenem Dritten (Stalnaker-Logik)
AL	Adams' System für indikativische Konditionale

Kapitel V (Parakonsistente Logik)

$\models_{\mathbf{D5}}$	diskussive Folgerung (Jaśkowski)
$\rightarrow_d$	diskussive Implikation
A°	Konsistenzoperator (da Costa)
A^n	n-fache ($n \geq 1$) Iteration von $^\circ$
Des	Menge designierter Werte
$\mathbf{C}_0, ..., \mathbf{C}_\omega$	da Costa-Systeme ($\mathbf{C}_0 = \mathbf{KL}$)
D5	= **J**, diskussive Logik
$\mathbf{D5}^d$	= $\mathbf{D}_2$, diskussive Logik mit $\rightarrow_d$
L3	Kleene-Logik
LP	Logik der Paradoxie (Priest)
$\mathbf{LP}^{\rightarrow}$	**LP** mit Implikation
RM3	dreiwertige Mingle-Logik

Kapitel VI (Relevanzlogik)

$\Vdash$	1. Sequenzenzeichen in **DML** ("first degree entailment") 2. Sequenzenzeichen in der Bündellogik
$(W,^*)$	Routley-Rahmen mit Routley-Stern *
$(W, 0, R)$	A-Rahmen, mit $0 \in W$ und dreistelliger Zugangsrelation R
$a \leq b$	Abkürzung für $R0ab$
R^2	aus R definierte vierstellige Relation

$(W, 0,^*, R)$	B- bzw. C-Rahmen (Routley-Meyer-Rahmen)
$\circ$	relevante Konjunktion (Fusion)
$+$	relevante Disjunktion (Fission)
,	extensionale Bündelung (von Prämissen)
;	intensionale Bündelung
DML	DeMorgan-Logik
N4	Priests vierwertige Logik mit nichtnormalen Welten
$\mathbf{N4}^{\rightarrow}$	implikatives Fragment von **N4**
$\mathbf{C}_{\rightarrow}$	Logik der A-Rahmen
$\mathbf{C}_{\neg\rightarrow}$	Logik der B-Rahmen (im $\{\neg, \rightarrow\}$-Fragment)
C	Logik der C-Rahmen
BCI	relevante kontraktionsfreie Implikationslogik
BCK	kontraktionsfreie Implikationslogik
DW	Grundsystem der Bündellogik
R	Standardsystem der Relevanzlogik
RM	Mingle-Logik

Kapitel VII (Anfechtbares Schließen)

$\vdash_K, \mathrm{Cn}_K$	materiales Folgern (unter Annahmen K)
$\vdash^+$	beliebige Erweiterung der Relation $\vdash$ in KL
$\mid\!\sim_K$, C_K	vorsichtiges materiales Folgern (mit Defaults K)
K_X	Menge der maximal mit X konistenten Teilmengen von K
s_K	Auswahlfunktion für K
$\leq_K$	Präferenzrelation für K
cwa_K	Prämissenwahl nach *Closed World Assumption*
$\mid\!\sim_{\kappa,<}$	materiale Folgerung in präferenziellen Modellen
$\vdash_R, \mathrm{Cn}_R$	materiales Folgern (unter Regeln R)
$\mathrm{mat}(R)$	Materialisierung der Regeln R
$\mid\!\sim_R, \mathrm{C}_R$	vorsichtiges materiales Folgern mit Default-Regeln
R_X	maximale Teilmengen der Regeln R, unter denen X konsistent abgeschlossen werden kann
$(A, \mathcal{B}, C)$	Default-Regel mit Antezedens A, Bedingungen $\mathcal{B}$ und Konklusion C
(D, X)	Default-Theorie mit Defaults D und Annahmen X
$\prec$	Wohlordnung von Defaults
$Anw(D, X)$	auf X anwendbare Defaults in D
$Erw(D, X)$	Erweiterungen von X durch Anwendung von Defaults in D
$\mid\!\sim_D, \mathrm{C}_D$	Folgern aus Defaults D

Kapitel VIII (Belief Revision)

$K + A$	Expansion von K mit A
$K - A$	Kontraktion von K um A
$K * A$	Revision von K mit A
$\Diamond$	ernsthafte Möglichkeit
$\sqsupset$	Ramsey-Konditional
$K \perp A$	Restefamilie
s_K	Auswahlfunktion für K
$\leq_K$	Präferenzrelation für K
$K \angle A$	minimale A-implizierende Teilmengen von K
$<$	relative Sicherheit
$\prec_{(K)}$	epistemische Verankerung (in K)
$\mathfrak{S}_K$	Sphärenmenge um K
S, S', S''	Sphären
$\mathfrak{S}_K/M$	Sphären um K, die M schneiden
C^*_K	durch Revision von K erzeugtes Folgern
$\mid\!\sim_\top$	parakonsistentes Folgern

4. Literaturhinweise

Für Philosophen gibt es ein großes Angebot an Einführungen und Übersichten (*Companions* und *Guides*), die über die elementare Darstellung der Aussagen- und Prädikatenlogik hinausgehen und sich in den Bereich der Philosophische Logik wagen. Zudem ist die *Stanford Encyclopedia of Philosophy* eine beinahe immer zuverlässige und bequem, d.h. *online* erreichbare Quelle. Es folgen einige Hinweise auf themenübergreifende Darstellungen, die besonders empfohlen werden können.

- J. van Benthem, *Modal Logics for Open Minds* [30]. Ein engagiert geschriebenes Buch, das auch neuere Entwicklungen beschreibt, wenn auch manchmal nur skizzenhaft. Fördert das Denken in Strukturen.
- P. Blackburn, *Modal Logic* [31]. Modernes, anspruchsvolles Lehrbuch.
- J.P. Burgess, *Philosophical Logic* [43]. Eine Einführung in eine interessante Themenauswahl. Bestechend flüssig zu lesen und dabei doch präzise.
- B.F. Chellas, *Modal Logic* [53]. Stellt die Modallogik recht zugänglich und vollständig bis etwa zum Ende der 70er Jahre dar.
- U. Friedrichsdorf, *Einführung in die klassische und intensionale Logik* [81]. Eine leider oft übersehene aber sehr empfehlenswerte Einführung. Sehr genau und umfassend.

- D.M. Gabbay & F. Guenthner, *Handbook of Philosophical Logic* [93]. In beiden Auflagen eine so gut wie vollständige Bestandsaufnahme zum jeweiligen Zeitpunkt. Beiträge von den jeweils besten Kennern ihres Gebiets. Man beachte, daß Handbuchartikel meist Resultate nennen, ohne die jeweiligen Beweise vorzustellen.
- D.M. Gabbay, C.J. Hogger & J.A. Robinson, *Handbook of Logic in Artificial Intelligence and Logic Programming* [94]. Philosophische Logik aus der Perspektive der Informatik.
- R. Goldblatt, *Logics of Time and Computation* [113]. Eine kurze und dabei erstaunlich umfassende Einführung in die Modallogik, einschließlich Dynamischer Logik.
- L. Humberstone, *The Connectives* [143]. Ein schier unerschöpflicher Fundus für ein weites Spektrum von Themen. Präzise und elegant.
- L. Humberstone, *Philosophical Applications of Modal Logic* [144]. Ditto.
- D. Jacquette, *A Companion to Philosophical Logic* [152]. Eine Sammlung von Beiträgen, die eine sehr breite Sicht auf Philosophische Logik insbesondere an der Schnittstelle zur Sprachphilosophie bietet.
- G. Priest, *An Introduction to Non-Classical Logic* [227]. Einführung in die Logik mit einem engagierten Blick auf Alternativen zur klassischen Logik. Der Autor bevorzugt konsequent Baumkalküle (Tableaux). (Auf Deutsch erhältlich.)
- S.L. Read, *Thinking About Logic* [241]. Eine Einführung mit wenig formalem Aufwand sowohl in die Philosophie der Logik als auch in die Philosophische Logik. (Auf Deutsch erhältlich.)

II.

ZEITLOGIK

> Time present and time past
> Are both perhaps present in time future
> And time future contained in time past.
> *T.S. Eliot*

1. Sein und Zeit

Von Augustinus stammt eine oft zitierte Bemerkung zu der Frage "Was ist die *Zeit*?" (*Confessiones*, XI 14):

> Wenn mich niemand darüber fragt, so weiß ich es; wenn ich es aber jemandem auf seine Frage erklären möchte, so weiß ich es nicht.

Eine ähnliche Beobachtung könnte man machen, wenn man danach fragt, was der *Raum* oder *Existenz* oder *Identität* sei. Das sind Begriffe, mit denen wir völlig vertraut umgehen. Wir wissen Vieles über Raum, Existenz und Identität, z.B., daß Deutschland östlich von England liegt, daß es keinen zweiten Erdmond gibt, oder daß der Abendstern identisch mit dem Morgenstern ist. Jedoch *was* das ist, von dem wir so vieles wissen: das wissen wir nicht – jedenfalls nicht ohne eine gehörige theoretische Anstrengung.

Aus der Not können wir versuchen, eine Tugend zu machen. Wir können versuchen, all das, was wir sicher über die Zeit wissen in einer Liste zusammentragen: zum Beispiel, daß alles Zukünftige einmal vergangen sein wird und alles Vergangene einmal zukünftig war, oder daß alles was in der Zukunft eines zukünftigen Zeitpunktes liegt, selbst zukünftig ist, u.s.w.. Je vollständiger diese Liste ausfiele, umso genauer würde sie festhalten, was Vergangenheit und Zukunft *an sich* ist, d.h. unabhängig davon, was in der Vergangenheit geschehen ist und in der Zukunft geschehen wird. Wenn wir so wüßten, was Vergangenheit bzw. Zukunft an sich ist, dann wären wir der Beantwortung von Augustinus' Frage offenbar schon ein gutes Stück näher gekommen. Vielleicht gibt es über "die Zeit" (an sich) gar nicht mehr zu wissen, als die ganz allgemeinen Eigenschaften von Vergangenheit und Zukunft. Genauer gesagt: Vielleicht ist die Rede von "der Zeit" nur eine irreführende Reifizierung, abgeleitet von einer Klassifizierung von Ereignissen als vergangen und zukünftig. Wenn das so ist, dann stellt Augustinus tatsächlich eine *logische* Frage: Wie läßt sich das Verhältnis von Vergangenheit und Zukunft (und vielleicht noch anderer zeitlicher Bestimmungen) ganz allgemein beschreiben? Es ist diese Frage jedenfalls, die sich die Zeitlogik stellt.

Man betrachte das folgende Beispielargument (aus [43, S. 13]):

(1) Publius und Quintus werden ihre Stimmen abgeben, jedoch nicht gleichzeitig.

Also:

(2) Entweder stimmt erst Publius und dann Quintus ab oder es stimmt erst Quintus und dann Publius ab.

Offensichtlich ist das ein guter Schluß. Es ist jedoch keiner, der sich in der klassischen Aussagenlogik wiedergeben läßt. Woran liegt das?

Schauen wir uns einen Teilsatz der Konklusion an:

(3) Publius stimmt ab (P) und dann stimmt Quintus ab (Q).

Können wir *und-dann* als eine Wahrheitsfunktion auffassen? Dann wäre die einfache Konjunktion $P \wedge Q$ für die Wahrheit von P *und dann* Q sicher notwendig aber möglicherweise nicht hinreichend. D.h. wir betrachten die Hypothese, daß P und dann Q soviel bedeutet wie $(P \wedge Q) \wedge C$, wobei C eine Bedingung darstellt, die wir hier nicht näher untersuchen müssen.

Angenommen nun, (3), d.h. $(P \wedge Q) \wedge C$ ist wahr. Dann ist $P \wedge Q$ und also $P \leftrightarrow Q$ wahr. Also ist $(P \wedge Q) \wedge C$ äquivalent zu $(Q \wedge Q) \wedge C$. Aber das bedeutet, daß wenn P *und-dann* Q wahr ist, dann ist auch Q *und-dann* Q wahr, d.h. Quintus stimmt zweimal ab – was wir als falsch unterstellen dürfen.

Das Ergebnis unserer Überlegung können wir so zusammenfassen: (2) ist in der klassischen Aussagenlogik nicht formalisierbar. Gleiches gilt für (1). Die klassische Aussagenlogik ist eine Logik des Seins: Wahrheit an einem Punkt, oder zeitlose Wahrheit. Wahrheit in der Zeit, d.h. den Wechsel von Wahrheitswerten über einen Raum von Zeitpunkten, kann sie nicht modellieren.

2. Zeit in der klassischen Prädikatenenlogik

In (1) und (2) kommen in keinem offensichtlichen Sinne Quantoren vor. So fallen diese Sätze nicht offensichtlich in den Modellierungsbereich dessen, was wir uns typischerweise unter Prädikatenlogik (PL) vorstellen. Aber man betrachte zum Vergleich den folgenden Satz:

(5) Deutschland liegt in Europa.

Auch in diesem Satz kommen – auf den ersten Blick – keine Quantoren vor. Seine Formalisierung in der PL ist dennoch naheliegend:

(6) *Liegt-in*(Deutschland, Europa).

Hier setzen wir voraus, daß wir in der Sprache ein zweistelliges Prädikat, *Liegt-in*, zur Verfügung haben. Nun könnte es sein, daß wir Länder als Mengen von geographischen Punkten und (6) als die Behauptung einer Mengeninklusion auffassen wollen. D.h. (6) ist dann einfach nur kurz für

(7) $\forall x(x \in \text{Deutschland} \rightarrow x \in \text{Europa})$.

So betrachtet, enthält (5) eine versteckte Quantifikation. Können wir mit (1) und (2) ähnlich verfahren und somit das Argument in der klassischen PL rekonstruieren? – Ja, das können wir.

Es seien x, y Variablen über Zeitpunkte. Mit x_0 wollen wir den Zeitpunkt bezeichnen, zu dem eine Formel auf ihre Wahrheit geprüft wird. Der Ausdruck $x < y$ bedeute, daß y in der Zukunft von x liegt; wir wollen das im strikten Sinne verstehen, so daß $<$ nicht reflexiv ist. Py (Qy) bedeute, daß Publius (Quintus) zum Zeitpunkt y abstimmt. Dann übersetzen wir wie folgt.

(1) Publius und Quintus werden ihre Stimmen abgeben, jedoch nicht gleichzeitig.

(1′) $\exists x(x_0 < x \wedge Px) \wedge \exists x(x_0 < x \wedge Qx) \wedge \neg\exists x(Px \wedge Qx)$.

(2) Entweder stimmt erst Publius und dann Quintus ab oder es stimmt erst Quintus und dann Publius ab.

(2′) $\exists x(x_0 < x \wedge Px \wedge \exists y(x < y \wedge Qy)) \vee \exists x(x_0 < x \wedge Qx \wedge \exists y(x < y \wedge Py))$.

Ein *Modell* $\mathcal{M} = (T, \prec, I)$ für eine Sprache mit den gerade verwendeten Ausdrücken sähe so aus:

- Eine nichtleere Menge T (Zeitpunkte, mit ausgezeichnetem Punkt 0);
- eine Relation $\prec \subseteq T \times T$ (zeitliche Abfolge);
- eine Interpretation I so, daß $I(P) \subseteq T$, $I(Q) \subseteq T$, und $I(<) = \prec$.
- Alles Weitere wie üblich in der klassischen PL.

Nun folgt (2′) aus (1′) genau dann, wenn jedes Modell, das (1′) erfüllt auch (2′) erfüllt. Das ist aber nicht der Fall, wie das folgende Modell mit drei Punkten $0, a, b$ zeigt:

$$\begin{array}{ccc} & & \bullet\, a\;\; P \\ & \nearrow & \\ 0\,\bullet & & \\ & \searrow & \\ & & \bullet\, b\;\; Q \end{array}$$

Das Modell erfüllt (1′) jedoch nicht (2′): Publius und Quintus geben in der Zukunft zu verschiedenen Zeitpunkten, a und b, ihre Stimmen ab. Jedoch liegt keiner dieser beiden Punkten jeweils in der Zukunft des anderen.

Als eine Darstellung der Zeit, ist dieses Modell natürlich überraschend. Es kommt dadurch zustanden, daß wir "vergessen" haben eine wichtige Bedingung zu erwähnen. Die Zeitpunkte stellen wir uns als linear miteinander verbunden vor, d.h. in Form einer *Kette*:

(Lin) $$a \prec b \text{ oder } a = b \text{ oder } b \prec a$$

Wenn wir nur solche Modelle betrachten, in denen die Zeitpunkte als Kette angeordnet sind, folgt dann (2′) aus (1′)? – Ja:

> Angenommen (1′) ist wahr am Punkt 0. Dann gibt es Punkte a und b mit $0 \prec a$, $0 \prec b$ und $a \neq b$ sowie $a : P$ und $b : Q$. Da $a \neq b$, so folgt aufgrund der Bedingung , daß entweder $a \prec b$ oder $b \prec a$. Im ersten Fall ist das erste Disjunkt von (2′) wahr, im zweiten Fall das zweite Disjunkt.

Wenn wir die Modelle so allgemein beschreiben, wie wir das zunächst getan haben, d.h. ohne die Kettenbedingung (Lin), dann ist der Schluß von (1) auf (2) nicht gültig; denn wir werden Modelle finden, in denen (1) wahr und (2) falsch ist. In jeder durch die Bedingung (Lin) eingeschränkten Klasse von Modellen, ist der Schluß jedoch gültig. Denn die Hinzunahme von (Lin) als Prämisse hat den Effekt, daß wir jetzt nur noch lineare Modelle der Zeit betrachten. In solchen Modellen folgt (2) aus (1). Damit haben wir eines der Kernthemen der Zeitlogik an einem Beispiel skizziert: Wir wollen herausfinden, welche Arten von Zeitstrukturen $(T, \prec)$ welche Folgerungen stützen, und umgekehrt, welche Folgerungen welche Zeitstrukturen erfordern.

3. Autonome Zeitlogik

Wir haben gesehen, daß die Prädikatenlogik durchaus geeignet ist, Zeitstrukturen zu beschreiben. Was (zeit)logisch aus Behauptungen über Zeit folgt, läßt sich im Prinzip im Rahmen der PL behandeln. Ist Zeitlogik also eine Abteilung der PL?

Im Prinzip kann man die Zeitlogik so betreiben. In PL-Sprachen kann man sogar mehr über Zeit sagen als in den aussagenlogischen Sprachen, die wir gleich betrachten werden. Dennoch gibt es Gründe, auch einen anderen, aussagenlogischen Ansatz zu verfolgen und diesen in der Regel sogar vorzuziehen:

· Insofern Formalisierung immer ein Mittel zu einem Zweck ist, sollte man keine Mittel wählen, die über den Zweck deutlich hinausgehen. Was uns an der Logik der Zeit interessiert, können wir, wie wir sehen werden, mit Mitteln beschreiben, die deutlich unterhalb der Komplexität von PL-Sprachen liegen.

· PL-Sprachen sind zwar sehr ausdrucksstark; das macht sie aber zugleich fragil. Wichtige logische Eigenschaften von formalen Systemen in einfacherer Sprache brechen zusammen, wenn die Sprache zur vollen PL angereichert wird. Insbesondere können wir von Zeitlogiken typischerweise zeigen, daß sie entscheidbar sind. Das gilt für PL bekanntlich nicht. So bilden Zeitlogiken in einer einfachen, aussagenlogischen Sprache entscheidbare Fragmente der PL ab.

· Typische Sprecher wollen mit dem Satz “Deutschland liegt in Europa” nicht etwas über das Verhältnis zweier Mengen aussagen. Ebensowenig geben (1′) und (2′) in irgendeinem naheliegenden Sinne die Struktur der Sätze (1) und (2) wieder. Wenn wir wissen möchten, wie Zeitsprache im Deutschen – einschließlich “Philosophischem Deutsch” – funktioniert, sind die Paraphrasen (1′) und (2′) dazu wenig geeignet.

Die Grundidee des autonomen Ansatzes beruht auf der Beobachtung, daß die Kopula in natürlichen Sprachen temporal ist (Präsens “ist”, Perfekt “war”, Futur “wird sein”, etc.).[1] Dabei kann ein grammatischer Tempus-Modus durchaus verschiedene Bedeutungen haben. So bezeichnet im Deutschen die Kopula “ist” im Präsens manchmal Gegenwartsprädikation (“Jetzt ist Schluß”) und manchmal atemporale Prädikation (“Schnee ist weiß”).

Im autonomen Ansatz wird nun das Tempus der Prädikation als Satzoperator herausgehoben und zum Gegenstand der logischen Untersuchung

[1] Die Kopula kann natürlich auch in einem nicht-temporalen Sinne *modal* sein. Einige dieser weiteren Modi des “Seins” werden uns später beschäftigen.

gemacht. Die folgende Liste enthält die in diesem Sinne naheliegendsten Zeitoperatoren:

$\mathsf{P}A$	in der Vergangenheit irgendwann: A
$\mathsf{F}A$	in der Zukunft irgendwann: A
$\mathsf{H}A$	in der Vergangenheit immer: A
$\mathsf{G}A$	in der Zukunft immer: A

Zunächst interessiert uns hier nur der Zukunftoperator F. Er funktioniert wie folgt. Die Wahrheit von Aussagen soll grundsätzlich an Zeitpunkten geprüft werden. Wenn t ein solcher Zeitpunkt ist, dann besagt $\mathsf{F}A$, daß A zu einem Zeitpunkt in der Zukunft von t wahr ist. Wenn wir so nach der Wahrheit von "Publius wird abstimmen" fragen, dann können wir das nur unter Bezugnahme auf einen Ausgangszeitpunkt t tun: Wir fragen, ob "Publius stimmt ab" wahr ist an einem Punkt in der Zukunft von t, d.h. wir fragen zum Zeitpunkt t, ob "F(Publius stimmt ab)" wahr ist. Der Operator F ist also wie ein Zeittransporter, der immer in eine Richtung fährt:

– Wir befinden uns an einem Zeitpunkt t.
– $\mathsf{F}A$ zum Zeitpunkt t behauptet, daß wir in der "Zukunftsrichtung" auf einen Zeitpunkt t' treffen, an dem A wahr ist.
– Wenn jetzt A (in $\mathsf{F}A$) wiederum von der Form $\mathsf{F}B$ sein sollte (oder auf eine solche Formel führt, wie im Fall (2*) unten), dann fahren wir vom Punkt t' weiter in derselben Richtung fort auf der Suche nach einem Punkt t'' an dem B wahr ist, u.s.w.

Allein mit Hilfe des F-Operators können wir so (1) und (2) in unserem Beispielargument durch die Formeln (1*) und (2*) wiedergeben:

(1) Publius und Quintus werden ihre Stimmen abgeben, jedoch nicht gleichzeitig.

(2) Entweder stimmt erst Publius und dann Quintus ab oder es stimmt erst Quintus und dann Publius ab.

(1*) $\mathsf{F}P \wedge \mathsf{F}Q \wedge \neg\mathsf{F}(P \wedge Q)$.

(2*) $\mathsf{F}(P \wedge \mathsf{F}Q) \vee \mathsf{F}(Q \wedge \mathsf{F}P)$.

Wenn der Übergang von (1) nach (2) ein guter Schluß ist, dann müssen wir uns jetzt um eine Interpretation des F-Operators bemühen so, daß sich der Übergang von (1*) zu (2*) als gültiger Schluß darstellt. Im nächsten Abschnitt machen wir dazu einen Anfang.

4. Kripke-Modelle (für $\mathcal{L}^{\mathsf{F}}$)

Es sei $\mathcal{L}$ eine Sprache der klassischen Aussagenlogik, d.h. die Formelmenge von $\mathcal{L}$ ist der Abschluß einer Menge ATM von atomaren Formeln (Atomen) unter einer funktional vollständigen Menge von Wahrheitsfunktionen; wir wählen hier $\{\neg, \wedge\}$. $\mathcal{L}^{\mathsf{F}}$ sei die Erweiterung dieser Sprache um den F-Operator. Mit FML wollen wir jetzt die Menge der Formeln von $\mathcal{L}^{\mathsf{F}}$ bezeichnen. (Später werden wir weitere Zeitoperatoren betrachten. Dann wird mit FML die Formelmenge der erweiterten Sprache gemeint sein. Im jeweiligen Kontext wird immer klar sein, welche Menge mit FML gemeint ist.) Die Menge FML sei also durch diese Bedingungen bestimmt:

1. ATM $\subseteq$ FML;
2. wenn $A, B \in$ FML, dann $\neg A, A \wedge B, \mathsf{F}A \in$ FML;
3. FML sei die kleinste Menge, welche die Bedingungen 1 und 2 erfüllt.

Eine "Zukunftssprache" $\mathcal{L}^{\mathsf{F}}$ interpretieren wir in Modellen

$$\mathcal{M} = (T, \prec, I).$$

Hier ist

- T eine nichtleere Menge (von Zeitpunkten);
- die Relation $\prec$ (früher als) verbindet Punkte in T, d.h. $\prec\, \subseteq T \times T$;
- die Interpretation $I : \text{ATM} \times T \longrightarrow \{0, 1\}$ bewertet Atome an Zeitpunkten als falsch (0) oder wahr (1).

Syntaktisch sind Junktoren Abbildungen von Formeln auf Formeln. In Modellen, d.h. semantisch, repräsentieren wir Formeln durch Mengen von Zeitpunkten: hier die Menge von Zeitpunkten zu denen die Formel wahr ist. So wird $\neg$ als eine Abbildung einer Teilmenge von T in deren Komplementärmenge interpretiert; das können wir der Klausel $(\neg)$ entnehmen. Mit F verfährt I nicht anders, das heißt

- I interpretiert die Operation F als eine Abbildung von Mengen von Zeitpunkten auf Mengen von Zeitpunkten, d.h. $I : \wp(T) \longrightarrow \wp(T)$. Welche Menge Wert der Abbildung ist, wird gleich aus der Klausel (F) hervorgehen.

Den Teil $\mathcal{K} = (T, \prec)$ nennt man eine *Struktur* (oder einen Rahmen, engl. *frame*), hier genauer eine *Kripke-Struktur*, und $\mathcal{M} = (\mathcal{K}, I)$ ist ein *Modell* für die Sprache $\mathcal{L}^{\mathsf{F}}$ auf der Struktur $\mathcal{K}$. Durch verschiedene Eigenschaften der Zeitordnung $\prec$ werden verschiedene Klassen von Strukturen festgelegt, in denen die Zeitoperatoren jeweils charakteristische Eigenschaften haben.

Das Ziel unserer Überlegungen ist eine Definition von Wahrheit in allen Modellen (logische Wahrheit) bzw. von Wahrheitsübertragung in allen Modellen (logische Folgerung). Dazu benötigen wir eine Definition der Wahrheit beliebiger Formeln an den Punkten beliebiger Modelle. Diese Definition baut auf der Festlegung der Wahrheitswerte der Atome auf, welche die Interpretation I leistet (s.o.). Wir bedienen uns dazu einer *Erfüllungsrelation* (hier vielleicht besser: Wahrmacherrelation), $\models_{\mathcal{M}} \subseteq T \times \mathrm{FML}$, die für jede Formel A und für jeden Punkt $a \in T$ (in einem Modell $\mathcal{M}$) bestimmt, ob A zum Zeitpunkt a wahr ist (in $\mathcal{M}$), d.h. wir lesen

$$a \models_{\mathcal{M}} A : a \text{ macht } A \text{ wahr in } \mathcal{M},$$

wobei wir im folgenden den Index $\mathcal{M}$ unterdrücken, wenn die Modellabhängigkeit im Kontext deutlich genug ist. Diese Erfüllungsrelation ist so definiert ("$\not\models$" ist kurz für "nicht $\models$"): Für alle $P \in \mathrm{ATM}$ und $A, B \in \mathrm{FML}$,

(P) $$a \models P \text{ gdw } I(P, a) = 1$$

$(\neg)$ $$a \models \neg A \text{ gdw } a \not\models A$$

$(\wedge)$ $$a \models A \wedge B \text{ gdw } a \models A \;\&\; a \models B$$

(F) $$a \models \mathsf{F}A \text{ gdw } \exists b : a \prec b \;\&\; b \models A$$

Definition 1. Sei $\mathcal{M}$ ein Modell auf einer Struktur $\mathcal{S} = (T, \prec)$ aus einer Klasse $\mathbb{S}$ von Strukturen.

1. Eine Formel A ist *wahr in* $\mathcal{M}$ gdw: A ist an allen Punkten in $\mathcal{M}$ wahr. (Notation: $\mathcal{M} \models A$.)
2. A ist *gültig* in der Strukturklasse $\mathbb{S}$ gdw: A ist in allen Modellen auf jeder Struktur $\mathcal{S} \in \mathbb{S}$ wahr. (Notation: $\mathbb{S} \models A$; die *Theorie* $Th(\mathbb{S})$ von $\mathbb{S}$ ist die Formelmenge $\{A : \mathbb{S} \models A\}$.)
3. A *folgt* aus $A_1, ..., A_n$ in $\mathcal{M}$ gdw: für alle Punkte $a \in T$, wenn $a \models A_1$ ... $a \models A_n$, dann $a \models A$. (Notation: $A_1, ..., A_n \models A$ in $\mathcal{M}$.)[2]
4. A *folgt* aus $A_1, ..., A_n$ in $\mathbb{S}$ gdw: A folgt aus $A_1, ..., A_n$ in allen Modellen auf allen Strukturen $\mathcal{S} \in \mathbb{S}$. (Notation: $A_1, ..., A_n \models A$ in $\mathbb{S}$.)

Im folgenden wird uns hauptsächlich Gültigkeit in bestimmten Strukturklassen, d.h. Wahrheit in allen Modellen auf Kripke-Strukturen einer bestimmten Art interessieren. Die allgemeinste Klasse von Strukturen, die

[2] Hier ist eine alternative Definition mit ganz anderem Resultat denkbar: ... für alle Modelle $\mathcal{M}$ gilt, wenn $\mathcal{M} \models A_1$ ... $\mathcal{M} \models A_n$, dann $\mathcal{M} \models A$. Die Wahl der einen oder anderen Definition ist vor allem entscheidend für den Status des Deduktionstheorems; vgl. dazu Kapitel III.6.

hier in Betracht kommt, ist die Menge *aller* Kripke-Strukturen, welche wir mit $\mathbb{K}$ notieren.

Eine im wesentlichen gleiche Überlegung wie wir sie oben bereits angestellt haben, zeigt, daß der Schluß von (1*) auf (2*) nicht gültig ist in der Klasse $\mathbb{K}$ aller Kripke-Strukturen. Da die Punkte nicht linear angeordnet sein müssen – d.h. so, wie wir uns die Zeit eigentlich vorstellen –, finden wir schnell ein Gegenmodell mit drei Punkten, in denen (1*) wahr und (2*) falsch ist (Übung!). Wenn der Operator F also die Bedeutung haben soll, die wir geben möchten, dann werden wir nur einen Teil der Menge aller Kripke-Strukturen betrachten wollen. Dennoch werden wir im nächsten Abschnitt kurz der Frage nachgehen, welche logischen Eigenschaften F in der Klasse *aller* Kripke-Strukturen hat. In gewisser Weise fragen wir so danach, was wir über den Zukunftsoperator F wissen, wenn wir nicht wissen – oder vorgeben nicht zu wissen –, ob die Zeitpunkte eine Kette bilden.

5. Die minimale "Zeitlogik" K

Die gerade gestellte Frage beantworten wir, indem wir ein axiomatisches System angeben, welches als Theoreme genau die Theorie $Th(\mathbb{K})$ der Kripke-Strukturen erzeugt, d.h. die Menge aller Sätze, die in beliebigen Modellen auf beliebigen Kripke-Strukturen wahr sind. Das folgende System **K** erfüllt die Bedingung.

Das System **K** (in F)

τ	Jede Tautologie
KF	$\neg \mathsf{F}A \wedge \mathsf{F}B \to \mathsf{F}(\neg A \wedge B)$
Regeln	$\dfrac{A \quad A \to B}{B}$ MP $\qquad \dfrac{\neg A}{\neg \mathsf{F}A}$ RNF

Die Axiome sind, wie im Eingangskapitel erklärt, im Sinne von Formelschemata zu verstehen; ebenso die Regeln. Es wird hier nicht lohnen, zwischen dem gerade definiertem formalen System und der Menge der Theoreme dieses Systems zu unterscheiden. Beide werden wir mit dem fetten Buchstaben **K** benennen. Diese Konvention fassen wir in einer streng genommen unsinnigen aber dennoch verständlichen Definition zusammen:

– **K** ist die Menge der Formeln, die aus den Axiomen und Regeln von **K** ableitbar sind.

- Mit **KL** bezeichnen wir die klassische Aussagenlogik, also die Menge der Tautologien.
- In dem Axiomenschema τ meinen wir die Menge der Tautologien in der erweiterten Sprache $\mathcal{L}^{\mathsf{F}}$. So ist z.B. $\mathsf{F}A \to \mathsf{F}A$ eine Tautologie (genauer: ein tautologisches Schema) im Sinne von τ, denn die Formel ist keine spezifisch zeitlogische Wahrheit sondern eine Einsetzungsinstanz von $B \to B$.

Unter der beabsichtigten Interpretation drückt KF etwas sehr Plausibles aus: Wenn nie A aber irgendwann einmal B wahr ist, dann ist irgendwann einmal A falsch und B wahr. Eine Variante von KF erhalten wir durch tautologische Umformung:

KF' $$\neg\mathsf{F}\neg(A \to B) \to (\mathsf{F}A \to \mathsf{F}B).$$

Um die Ableitungsregel RNF richtig zu verstehen, müssen wir an dieser Stelle den Begriff der Ableitung näher betrachten.

Wie in Kapitel I.2 erklärt, ist eine *Ableitung* (ein Beweis) einer Formel A in einem axiomatischen System S ist eine Folge $\langle A_0, ..., A_n\,[= A]\rangle$ von Formeln so, daß jede Formel A_i $(0 \leq i \leq n)$ entweder (1) ein Axiom von S ist oder (2) Konklusion aus dem A_i vorangehenden Formeln augrund einer der Regeln von S ist. Dieser Begriff der Ableitung (von Theoremen) ist ein Grenzfall des allgemeineren Begriffs der *Ableitung aus Annahmen.* Letzterer fügt den Bedingungen (1) und zwei eine weiter Möglichkeit hinzu: (3) A_i ist in der Menge der Annahmen enthalten. Die Ableitung eines Theorems ist in diesem Sinne eine Ableitung aus der leeren Annahmenmenge. Während die Notation $\vdash A$ festhält, daß A ein Theorem des betrachteten axiomatischen Systems ist, besagt $X \vdash A$, daß A unter möglichem Rückgriff auf die Annahmen in X abgeleitet werden kann. Wir nennen $X \vdash A$ auch eine Sequenz (des Systems). In der klassischen Aussagenlogik gibt es eine enge Beziehung zwischen Theoremen und Sequenzen, nämlich

DED. $$X, A \vdash B \text{ gdw } X \vdash A \to B,$$

Von rechts nach links gelesen, ist die Äquivalenz eine unmittelbare Folge von Modus Ponens, in umgekehrter Richtung handelt es sich um das *Deduktionstheorem.*

Regeln wie RNF sind ein wesentlicher Bestandteil der Zeitlogik, wie überhaupt der Modallogik. Leider führen diese modalen Regeln dazu, daß die Äquivalenz DED bricht. Denn die Folge $\langle\neg A, \neg\mathsf{F}A\rangle$ stellt im Sinne der Definition natürlich eine Ableitung von $\neg\mathsf{F}A$ aus der Annahme $\neg A$ dar, was wir so festhalten können:

(*) $$\neg A \vdash \neg\mathsf{F}A.$$

Aber wenn wir aus dieser Sequenz aufgrund von DED auf

(†) $$\neg A \to \neg \mathsf{F} A$$

schließen wollten, dann würden wir auf eine Formel schließen, die *kein* Theorem von **K** ist. Das ist sicher auch gut so, denn (†) drückt ja etwas offensichtlich Falsches aus: Was (zu einem Zeitpunkt) falsch ist, kann (später) nie wahr werden. Wie zu erwarten, ist die Formel (†) nach Definition 1 auch nicht gültig (in $\mathbb{K}$). Schließlich können wir noch beobachten, daß nach derselben Definition (∗) auch keine gültige Folgerung darstellt. Im Falle von **K** (und modalen Systemen im allgemeinen) passt also der Begriff einer Ableitung aus Annahmen ebensowenig zum Deduktionstheorem wie zum Begriff der Folgerung in Modellen.

Es gibt mehrere Möglichkeiten hier Abhilfe zu schaffen. Im Kapitel III.6 werden wir die Optionen eingehend besprechen. Hier werden wir uns damit behelfen, eine zweite, mit $\vdash$ verwandte Relation $\triangleright$ einführen, die gewissermaßen die Simulation einer Ableitung durch eine Formel anzeigt. Als Definition dieser Relation nehmen wir einfach die Äquivalenz DED und führen eine neue Notation ein:

(Def. $\triangleright$) $$A_1, ..., A_n \triangleright B \text{ gdw } \vdash A_1 \wedge \cdots \wedge A_n \to B \ (n \geq 0).$$

Jetzt gilt das Deduktionstheorem für $\triangleright$ – jedenfalls für endliche Annahmenmengen – *per definitionem*. Zwar ist $\neg \mathsf{F} A$ aus $\neg A$ im üblichen Sinne ableitbar – und also ist $\neg A \vdash \neg \mathsf{F} A$ eine richtige Sequenz –, aber $\neg A \to \neg \mathsf{F} A$ ist kein Theorem und also ist die Sequenz $\neg A \triangleright \neg \mathsf{F} A$ falsch.

Nach diesem längeren aber nötigen Exkurs über Ableitbarkeit in der Zeitlogik wenden wir uns jetzt wieder der Frage zu, wie die **K**-Regel RNF zu verstehen sei. Wie alle Ableitungsregeln hat sie die Aufgabe aus Theoremen als Prämissen Theoreme als Konklusionen zu produzieren. Den Satz 2 vorwegnehmend, wissen wir, daß Theoreme logische Wahrheiten sind. Unschwer läßt sich nachprüfen, daß die Regel zwingend von logischen Wahrheiten zu logischen Wahrheiten führt. Denn wenn $\neg A$ logisch, d.h. an allen Punkten wahr ist, dann wird an keinem Punkt künftig A wahr sein, d.h. $\neg \mathsf{F} A$ ist logisch wahr.

In Ableitungen *aus Annahme*n hat RNF jedoch nicht immer die Eigenschaft aus einer (kontingent) wahren Prämisse (z.B. $\neg P$, für ein Atom P) auf eine wahre Konklusion (hier: $\neg \mathsf{F} P$) zu führen. Das ist *per se* nicht schlimm und zeigt nur, daß Ableitbarkeit nicht ohne weiteres mit logischer Folgerung gleichzusetzen ist. Schlimm – weil falsch – wäre es, in einer solchen Situation nach DED auf $\neg P \to \neg \mathsf{F} P$ zu schließen.

Für eine ganze Klasse von Formeln als Prämissen der Regel RNF garantiert diese jedoch auch die Wahrheitsübertragung, wenn im Laufe einer Ableitung auf kontingente Annahmen zurückgegriffen wird. Sei die Prämisse zum Beispiel von der Form $\neg\mathsf{F}A$. Wenn diese Formel an einem Punkt a wahr ist, dann schließt die Zukunft von a Punkte mit A aus. Jeder Punkt a' in der Zukunft von a hat seine eigene Zukunft, die ein Teil der Zukunft von a ist. Wenn also die a-Zukunft keine A-Punkte enthält, dann gilt das auch für jeden Teil der a-Zukunft. D.h. unter der Annahme muß am Punkte a die Formel $\neg\mathsf{FF}A$ wahr sein. In allen Fällen, in denen die die Regel RNF die Wahrheit an einem Punkt von der Prämisse auf die Konklusion überträgt, darf dann auch konditionalisiert werden; im gerade behandelten Fall dürfen wir auf $a \models \neg\mathsf{F}A \to \neg\mathsf{FF}A$ schließen. In solchen Fällen gilt also das Deduktionstheorem.

Damit wollen wir die Erläuterungen zu den Axiomen und Regeln des Systems **K** beenden und kommen nun zu dem zentralen Resultat, welches die Axiomatik **K** und die Strukturklasse $\mathbb{K}$ zueinander in Beziehung setzt.

Satz 2. (Richtigkeit und Vollständigkeit)

1. *Jedes Theorem von* **K** *ist gültig in der Klasse* $\mathbb{K}$ *der Kripke-Strukturen, d.h.* $\mathbf{K} \subseteq Th(\mathbb{K})$ *(Richtigkeit im Hinblick auf* $Th(\mathbb{K})$*).*
2. *Jede in* $\mathbb{K}$ *gültige Formel ist in* **K** *ableitbar,*
   d.h. $Th(\mathbb{K}) \subseteq \mathbf{K}$ *(Vollständigkeit im Hinblick auf* $Th(\mathbb{K})$*).*
3. $\triangleright$*-Ableitbarkeit in* **K** *stimmt mit Folgerung in* $\mathbb{K}$ *überein,*
   d.h. $(\triangleright \text{ in } \mathbf{K}) = (\models \text{ in } \mathbb{K})$
   (Richtigkeit und Vollständigkeit von $\triangleright$ *im Hinblick auf* $\models$*).*

Wir werden den Satz erst im Kapitel III über Modallogik beweisen und uns an dieser Stelle mit einer Beweisskizze für 1 und 2 begnügen. (Teil 3 des Satzes folgt unmittelbar aus 1 und 2 aufgrund der Definition von $\triangleright$.) Für die Richtigkeit von **K** bezüglich $Th(\mathbb{K})$ nehme man an, eine Formel A sei ableitbar und zeige dann durch Induktion über die Länge einer Ableitung von A (siehe Kapitel I.2), daß A an einem beliebigen Punkt in einem beliebigen Kripke-Modell wahr wird. Für die Vollständigkeit, betrachte man eine beliebige *nicht* ableitbare Formel A und konstruiere ein Modell, in dem A falsch ist. Dann gilt umgekehrt: Wenn es für eine Formel kein Gegenmodell gibt, d.h. die Formel in $\mathbb{K}$ gültig ist, dann ist die Formel ableitbar. Zur Konstruktion von Gegenmodellen allein aufgrund der Information, daß eine Formel nicht ableitbar ist, werden wir uns der Methode nach Lindenbaum und Henkin bedienen.[3]

[3] Für die Modalllogik wurde die Methode von Makinson adaptiert; siehe [192].

Wir kehren kurz zu unserem Eingangsbeispiel über Publius und Quintus zurück. Können wir den beabsichtigten Schluß in der Logik **K** rekonstruieren? Dazu stehen uns nun nach Satz 2 grundsätzlich zwei Methoden zur Verfügung. Wir können versuchen, aus der Prämisse

$$\mathsf{F}P \wedge \mathsf{F}Q \;\wedge\; \neg\mathsf{F}(P \wedge Q) \tag{1}$$

die Konklusion

$$\mathsf{F}(P \wedge \mathsf{F}Q) \;\vee\; \mathsf{F}(Q \wedge \mathsf{F}P) \tag{2}$$

in **K** abzuleiten. Oder wir können versuchen nachzuweisen, daß in jedem Modell die Sequenz (1) $\triangleright$ (2) wahr ist. Wir haben oben schon angedeutet, daß letzterer Versuch scheitern wird. Wir wollen uns den Versuch dennoch einmal ansehen, um zu beobachten, woran er letztlich scheitert.

Dazu betrachten wir einen beliebigen Punkt a in einem beliebiges Modell und nehmen an, daß (1) an diesem Punkt wahr ist. Dann gibt es Punkte b und c in der Zukunft von a derart, daß

$$b \models P \text{ und } c \models Q, \text{ und } \neg\exists x \succ a : x \models P \wedge Q. \tag{1}$$

Die dritte Bedingung in (1) schließt insbesondere aus, daß $b = c$. Das kleinste Modell, das eine solche Situation darstellt, besteht aus genau drei Punkten:

A.
$$\begin{tikzcd} & \bullet\, b\ P \\ a\,\bullet \arrow[ur] \arrow[dr] & \\ & \bullet\, c\ Q \end{tikzcd}$$

Hier soll der Pfeil für die Relation $\prec$ stehen, also $x \longrightarrow y = x \prec y$. Die beabsichtigte Konklusion (2) behauptet, daß am Punkt a auch die Formel (2) wahr ist, d.h.

$$\exists x \succ a : x \models P \wedge \mathsf{F}Q \text{ oder } \exists y \succ a : y \models Q \wedge \mathsf{F}P.$$

Nun gib es tatsächlich einen Punkt in der Zukunft von a, nämlich b, an dem P wahr ist – aber von dort verweist kein Pfeil auf einen Q-Punkt, d.h. $\mathsf{F}Q$ ist am Punkt b falsch. Gleiches gilt für $Q \wedge \mathsf{F}P$. Abhilfe würde die folgende Bedingung schaffen:

$$a \prec b \;\&\; a \prec c \;\;\Rightarrow\;\; b \prec c \text{ oder } c \prec b \text{ oder } b = c. \tag{$*$}$$

Wir wissen, daß die Möglichkeit $b = c$ durch unsere Annahme (1) ausgeschlossen ist. Die Bedingung würde also erzwingen, daß wir das Bild A in der einen oder anderen Weise vervollständigen:

$$\text{B.}\qquad \begin{array}{ccc} & & \bullet\ \ b\,P \\ & \nearrow & \\ a\,\bullet & & \downarrow \\ & \searrow & \\ & & \bullet\ \ c\,Q \end{array} \qquad\qquad \begin{array}{ccc} & & \bullet\ \ b\,P \\ & \nearrow & \\ a\,\bullet & & \uparrow \\ & \searrow & \\ & & \bullet\ \ c\,Q \end{array}$$

Jetzt verweist ein Pfeil von b auf einen Q-Punkt bzw. von c auf einen P-Punkt und somit ist die Formel (2) wahr am Punkt a!

Mit $(*)$ haben wir eine Bedingung für Kripke-Strukturen identifiziert, welche die Gültigkeit des Schlusses von (1) auf (2) garantiert. Es stellen sich zwei naheliegende Fragen:

1. Ist $(*)$ die schwächste, d.h. nicht nur hinreichende sondern auch notwendige Struktur-Bedingung, welche den Schluß garantiert?
2. Gibt es eine Formel, welche diese Bedingung "ausdrückt"? Hier fragen wir nach einer Formel A so, daß (a) A in allen Kripke-Strukturen unter Bedingung $(*)$ gültig ist, und (b) wenn A in einer Struktur gültig ist, dann erfüllt diese Struktur die Bedingung $(*)$.

Die zweite Frage löst die Suche nach einer *Korrespondenz* aus zwischen einem Axiomenschema A einerseits und der Strukturbedingung (*) andererseits so, daß

$$\mathbf{K} + A = Th(\mathbb{K} + (*)).$$

Das System $\mathbf{K} + A$ wäre dann die minimale Logik in der sich der Schluß von (1) auf (2) als Sequenz $(1) \triangleright (2)$ reproduzieren läßt. Fragen solcher Art werden uns im Abschnitt 8 unter dem Stichwort "Korrespondenzen" beschäftigen. Dort werden wir zeigen, daß die erste Frage zu bejahen (p. 57) und die zweite zu verneinen (p. 54) ist.

6. "Multimodale" Zeitlogik

Wir schauen nicht nur vorwärts in die Zukunft, sondern auch zurück in die Vergangenheit. So wollen wir zum Beispiel sagen können

(4) Wenn Publius bisher nicht abgestimmt hat, dann wird er noch abstimmen.

Hier sind offenbar zwei Operationen auf dem Satz

P Publius stimmt ab

am Werk. Den Zukunftsoperator F kennen wir schon. Es fehlt noch ein sich dazu spiegelbildlich verhaltender Vergangenheitsoperator P. Dann können wir (4) so formalisieren:

(4*) $\neg\mathsf{P}P \wedge \neg P \to \mathsf{F}P$

Den Operator P wollen wir, wie zuvor schon F, in Modellen interpretieren. Den einstelligen Operator F haben wir mit einer zweistelligen Relation $\prec$ auf der Grundmenge T modelliert. Genauso können wir auch mit P verfahren. Wir gehen wieder aus von einer Struktur $(T, \succ)$ und beziehen die P-Sprache darauf mit Hilfe einer Interpretation I. Im weiteren legen wir die Wahrheitsbedingung für P-Formeln so fest:

$$(\mathsf{P}) \qquad a \models \mathsf{P}A \text{ gdw } \exists b : a \succ b \ \&\ b \models A.$$

Ein kurzer Vergleich mit der Wahrheitsbedingung für (F),

$$(\mathsf{F}) \qquad a \models \mathsf{F}A \text{ gdw } \exists b : a \prec b \ \&\ b \models A,$$

zeigt, daß wir hier eigentlich nur Zeichen ausgetauscht haben: P gegen F und $\succ$ gegen $\prec$. Wenn also die Logik **K** für F zu Strukturen des Typs $(T, \prec)$ im Sinne des Satzes 2 paßt, dann paßt sicher der Strukturtyp $(T, \succ)$ zur Logik **K** für P, in der wir die F-Schemata austauschen gegen KP und RNP, also

$$(\neg\mathsf{P}A \wedge \mathsf{P}B) \to \mathsf{P}(\neg A \wedge B) \quad \text{und} \quad \frac{\neg A}{\neg\mathsf{P}A}.$$

Der Beweis, daß **K** für P zum Strukturtyp $(T, \succ)$ paßt, ist wenig mehr als eine Kopie des Beweises für Satz 2.

Im nächsten Schritt wollen wir eine Sprache interpretieren, in der beide Operatoren vorkommen. Dazu können wir die Kripke-Strukturen $(T, \succ)$ und $(T, \prec)$ zu einer Struktur $(T, \succ, \prec)$ zusammenfügen, und geben dann die Wahrheitsbedingungen für F und P wie zuvor an. Beispielsweise sieht dann die Wahrheitsbedingung für die Formel (4*) in einem Modell $(T, \succ, \prec, I)$ so aus:

$$\text{Wenn } a \not\models P \text{ und } \neg\exists x : a \succ x \text{ und } x \models P, \text{ dann } \exists x : a \prec x \text{ und } x \models P.$$

Wie schon die Struktur $(T, \succ, \prec)$, so erhalten wir die dazu passende Logik durch Zusammensetzung: Wir kombinieren einfach die Logik **K** für F-Sprachen mit der Logik **K** für P-Sprachen. Kaum überraschend, läßt sich diese Logik so axiomatisieren:

Das System $\mathbf{K}_0$ (in FP)

τ	Jede Tautologie
KF	$\neg\mathsf{F}A \wedge \mathsf{F}B \rightarrow \mathsf{F}(\neg A \wedge B)$
KP	$\neg\mathsf{P}A \wedge \mathsf{P}B \rightarrow \mathsf{P}(\neg A \wedge B)$
Regeln	$\dfrac{A \quad A \rightarrow B}{B}$ MP $\qquad \dfrac{\neg A}{\neg\mathsf{F}A}$ RNF $\qquad \dfrac{\neg A}{\neg\mathsf{P}A}$ RNP

Wir haben nun in unserer Sprache *zwei* Zeitoperatoren. Offensichtlich läßt sich der eine nicht auf den anderen durch eine Definition reduzieren: Was vergangen ist, läßt sich nicht aus dem, was in der Zukunft liegt, erklären – und ebensowenig umgekehrt. Diese beiden Operatoren haben wir anhand zweier Relationen modelliert. Auch diese Relationen sind unabhängig voneinander. Es gibt keine Strukturbedingungen, welche die beiden Relationen zueinander in Beziehung setzen, also insbesondere keine aus denen die Reduzierbarkeit der einen auf die andere folgen könnten. So haben wir ein einfaches Beispiel für eine genuin *multimodale* – in diesem Fall bimodale – Logik.

Aber dieser Stand der Dinge ist für unsere Absicht, eine Logik der Zeit – jedenfalls im Alltagsverständnis – zu entwerfen, unbefriedigend. Damit die Relationen $\succ$ und $\prec$ den Fluß der Zeit in jeweils einer von zwei Richtungen darstellen können, wünschen wir uns ein sehr starkes Zusammenspiel zwischen beiden, nämlich das die eine das Spiegelbild der anderen sei. (Im Abschnitt 12 werden wir einen Gegenentwurf zur Spiegelbildlichkeit von Vergangenheit und Zukunft behandeln.) Dieses Verhältnis deuten die verwendeten Zeichen an; aber die Form der Zeichen ersetzt keine Festlegung ihrer Bedeutungen. Die inhaltliche Festlegung, welche die gewählten Zeichen erwarten läßt, ist diese:

(Spiegelung) $\qquad a \prec b$ gdw $b \succ a.$

In Kripke-Strukturen, welche diese Spiegelungsbedingung erfüllen, stehen die Relationen nicht mehr unvermittelt nebeneinander. So erhalten nun Formeln den Status logischer Wahrheit, in denen Vergangenheits- und Zukunftsoperatoren in ein Verhältnis zueinander gesetzt werden. Die folgenden zwei Schemata werden wir gleich als Axiome verwenden:

B. $\qquad A \rightarrow \neg\mathsf{P}\neg\mathsf{F}A \qquad A \rightarrow \neg\mathsf{F}\neg\mathsf{P}A.$

Wenn A wahr ist, dann war es immer schon wahr, daß A einmal wahr wird, und ebenso wird es immer wahr sein, daß A einmal wahr gewesen ist. Wir betrachten jetzt nur das linke Schema. Die folgenden zwei Beobachtungen sind für das rechte Schema völlig analog. Daß (Spiegelung) das Schema B (links) zu einer logischen Wahrheit macht, zeigen wir so:

(a) Es sei a ein beliebiger Punkt mit $a \models A$. Zu zeigen: $a \models \neg\mathsf{P}\neg\mathsf{F}A$, d.h. wenn $a \succ b$, dann gibt es ein x mit $b \prec x$ und $x \models A$. Wir nehmen daher ferner an, daß $a \succ b$ und also (Spiegelung!) $b \prec a$. Dann ist a der gesuchte Punkt x mit $b \prec x$ und $x \models A$.

Wir können uns auch umgekehrt davon überzeugen, daß in Kripke-Strukturen ohne (Spiegelung), das Schema B ungültig sein muß. Dazu genügt es, ein Modell zu betrachten, das die Bedingung nicht erfüllt, und in diesem Modell eine Instanz von B (z.B. ein Atom P für A) an einem Punkte falsch zu machen.

(b) Aus der Wahrheitsbedingung für $P \to \neg\mathsf{P}\neg\mathsf{F}P$ ergibt sich die Beschreibung eines falsifizierenden Punktes a ín einem Modell:

$$(*) \qquad a \models P \;\&\; \exists a \succ x \;\&\; x \not\models \mathsf{F}P \text{ (d.h. } \forall y : x \prec y \Rightarrow y \not\models P).$$

Man betrachte also ein Modell mit zwei Punkten a und b so, daß $a \models P$ und $a \succ b$ (und nicht $b \prec a$). Die Bedingung $(*)$ ist erfüllt.

Mit (a) und (b) haben wir tatsächlich gezeigt, daß jedes der zwei Schemata B in allen Kripke-Strukturen mit (Spiegelung) und nur in diesen gültig ist. D.h. jedes der Schemata B *korrespondiert* mit der Bedingung (im Sinne der Definition 9, p. 53).

Die einfachste Weise, das natürliche Zusammenspiel von Vergangenheit und Zukunft in den zu betrachtenden Strukturen zu verwirklichen, besteht darin, sie zu vereinfachen. Wir wählen eine der beiden Relationen aus. (Die andere können wir per (Spiegelung) definieren, falls gewünscht.) Sei $\prec$ die gewählte Relation. Dann betrachten wir Strukturen von Typ $(T, \prec)$ und interpretieren darin **P**-Formeln so:

$$(\mathsf{P}) \qquad a \models \mathsf{P}A \text{ gdw } \exists b : b \prec a \;\&\; b \models A.$$

Nun erst haben **P** und **F** ihre beabsichtigte Bedeutung: Der eine Operator blickt in die eine, der andere in die andere Richtung derselben Relation.

Wie wir oben gesehen haben drückt sich das durch (Spiegelung) bewirkte Zusammenspiel von **P** und **F** in den Formelschemata B aus. Diese werden etwas übersichtlicher, wenn wir sie mit Hilfe zweier neuer Operatoren formulieren, die wir auf S. 34 schon kurz erwähnt haben.

$$(\mathsf{H}) \qquad a \models \mathsf{H}A \text{ gdw } \forall b : b \prec a \;\Rightarrow\; b \models A$$

$$(\mathsf{G}) \qquad a \models \mathsf{G}A \text{ gdw } \forall b : a \prec b \;\Rightarrow\; b \models A$$

Offensichtlich besagt $\mathsf{H}A$ soviel wie: *A war immer der Fall*, während $\mathsf{G}A$ besagt: *A wird immer der Fall sein.* Und genauso wie $\exists$ und $\forall$ mittels der Negation gegenseitig definierbar sind, so verhält es sich auch mit dem existenzquantifizierenden P und dem allquantifizierenden H einerseits sowie F und G andererseits:

$$\mathsf{P} \equiv \neg\mathsf{H}\neg \quad \text{und} \quad \mathsf{F} \equiv \neg\mathsf{G}\neg\,.$$

Sprachen mit einem funktional vollständigen Paar dieser Operatoren – also jeweils einem Vergangenheits- und einem Zukunftsoperator – wollen wir hier *Zeitsprachen* nennen. Interpretationen auf Kripke-Strukturen von Zeitsprachen, nennen wir in diesem Kapitel *Zeitmodelle*.

- Wenn $(T, \prec)$ also eine Kripke-Struktur ist, dann ist $(T, \prec, I)$ mit den Bedingungen (H) und (G) ein *Zeitmodell* auf $(T, \prec)$. Die Menge $Th_t(\mathbb{K})$ bezeichnet die Menge aller Formeln, die in beliebigen Zeitmodellen auf Kripke-Strukturen gültig sind – die Zeittheorie von $\mathbb{K}$.

Es sind H ("immer schon") und G ("künftig immer") nicht nur sehr natürliche Zeitoperatoren, sie eignen sich auch gut für eine unmittelbar einleuchende Formulierung von B und damit für die folgende axiomatische Darstellung von $Th_t(\mathbb{K})$.

Das System $\mathbf{K}_t$ (in GH)

Def.	$\mathsf{F} := \neg\mathsf{G}\neg$	$\mathsf{P} := \neg\mathsf{H}\neg$
τ	Jede Tautologie	
K	$\mathsf{H}(A \to B) \to (\mathsf{H}A \to \mathsf{H}B)$	$\mathsf{G}(A \to B) \to (\mathsf{G}A \to \mathsf{G}B)$
B	$A \to \mathsf{HF}A$	$A \to \mathsf{GP}A$
RN	$\dfrac{A}{\mathsf{H}A}$	$\dfrac{A}{\mathsf{G}A}$
MP	$\dfrac{A \quad A \to B}{B}$	

Das Schema B besagt nun: Was wahr ist war immer schon künftig einmal wahr und wird immer einmal wahr gewesen sein. Zusammengenommen, besagt RN, daß logische Wahrheiten immer ("ewig") wahr sind.

SATZ 3. *Eine Formel A einer Zeitsprache ist genau dann ein Theorem von $\mathbf{K}_t$, wenn $A \in Th_t(\mathbb{K})$, d.h. wenn A gültig ist in der Klasse aller Zeitmodelle auf beliebigen Kripke-Strukturen.*

Das System $\mathbf{K}_t$ gibt also die minimale Logik einer Sprache an, in der wir *echte* Vergangenheits- und Zukunftoperatoren haben, d.h. solche, die sich spiegelbildlich zueinander verhalten.

7. Einige Eigenschaften von $\mathbf{K}_t$

Wir betrachten jetzt einige Eigenschaften der Theoremmenge von $\mathbf{K}_t$. Einige dieser Eigenschaften gelten schon für die klassische Aussagenlogik **KL** und bleiben bei jeder Erweiterung $\mathbf{KL}^+$ von **KL** erhalten, also auch für $\mathbf{K}_t$ und mögliche Erweiterungen $\mathbf{K}_t^+$.

- Unter den *Erweiterungen* $\mathbf{L}^+$ einer Logik **L** wollen wir hier nur *schematische* Erweiterungen verstehen. Diese enstehen durch Hinzunahme weiterer Axiomen*schemata* bzw. *schematischer* Regeln, also aller Formeln bzw. Regeln einer bestimmten Gestalt. Unter der – hier immer zutreffenden – Annahme, daß **L** unter *Einsetzung* abgeschlossen ist, wird so auch jede Erweiterung $\mathbf{L}^+$ unter Einsetzung abgeschlossen sein.

Die erste Beobachtung besagt, daß eine Erweiterung von **KL** nicht nur den Theorembestand von **KL** sondern auch den Bestand an Ableitungen (Sequenzen) erweitert. Das folgt unmittelbar aus den Definitionen von "Ableitung" und "Erweiterung". Ableitungen in $\mathbf{K}^+$ greifen mindestens auf die Ressourcen (Theoreme) der klassischen Aussagenlogik **KL** zurück.

SATZ 4. (Klassische Konsequenz)
Wenn $A_1, ..., A_n \vdash A$ in **KL**, *dann $A_1, ..., A_n \vdash A$ in allen* $\mathbf{KL}^+$.

BEWEIS. Angenommen $A_1, ..., A_n \vdash A$ in **KL**. In **KL** – möglicherweise aber nicht in $\mathbf{KL}^+$! – gilt das Deduktionstheorem, weshalb wir durch fortgesetzte Anwendung $\vdash A_1 \rightarrow (A_2 \rightarrow ...(A_n \rightarrow A)...)$ in $\mathbf{KL} \subseteq \mathbf{KL}^+$ erhalten. Aus der Annahme A_1 folgt dann in $\mathbf{KL}^+$ per MP: $A_2 \rightarrow (A_3 \rightarrow ...(A_n \rightarrow A)...)$ ... und schließlich erhalten wir so die gewünschte Sequenz $A_1, ..., A_n \vdash A$ in $\mathbf{KL}^+$. ∎

Zweitens können wir feststellen, daß wenn B aus A logisch folgt (also $A \triangleright B$, d.h. $\vdash A \rightarrow B$), dann folgt auch jede Temporalisierung (Tempus) T von B aus der gleichen Temporalisierung von A.

- Unter einem *Tempus T* verstehen wir hier jede (endliche) Verkettung der Operatoren $\mathsf{G}, \mathsf{F}, \mathsf{H}$ und P.

Satz 5. (Monotonie)
Die Regeln RMT, $\dfrac{A \to B}{TA \to TB}$, *sind zulässig in allen* $\mathbf{K}_t^+$.

Beweis. Es genügt, den Satz für die einfachen Tempora $\mathsf{G}, \mathsf{F}, \mathsf{H}$ und P zu beweisen. Für zusammengesetzte Tempora folgt der Satz dann durch wiederholte Anwendung. Hier ist der Beweis für G und F (der für H und P geht genauso):

1	(1)	$A \to B$	Annahme
1	(2)	$\mathsf{G}(A \to B)$	(1), RNG
	(3)	$\mathsf{G}(A \to B) \to (\mathsf{G}A \to \mathsf{G}B))$	KG
1	(4)	$\mathsf{G}A \to \mathsf{G}B$	(3-4), MP

1	(1)	$A \to B$	Annahme
1	(2)	$\neg B \to \neg A$	(1), **KL**
1	(3)	$\mathsf{G}(\neg B \to \neg A)$	(2), RNG
	(4)	$\mathsf{G}(\neg B \to \neg A) \to (\mathsf{G}\neg B \to \mathsf{G}\neg A)$	KG
1	(5)	$\mathsf{G}\neg B \to \mathsf{G}\neg A$	(3-4), MP
1	(6)	$\neg\mathsf{G}\neg A \to \neg\mathsf{G}\neg B$	(5), **KL**
1	(7)	$\mathsf{F}A \to \mathsf{F}B$	(6), Def.

■

Drittens beobachten wir, daß in jeder Erweiterung von **K** die logische Äuivalenz von Formeln erhalten bleibt, wenn darin logisch äquivalente Formeln gegeneinander ausgetauscht werden. Diesen Satz werden wir im Folgenden oft anwenden, ohne auf die Anwendung eigens hinzuweisen.

Satz 6. (Ersetzbarkeit äquivalenter Formeln)
In allen Erweiterungen $\mathbf{K}_t^+$ *von* $\mathbf{K}_t$ *gilt: Wenn* $\vdash B \leftrightarrow B'$, *dann* $\vdash A \leftrightarrow A(B'/B)$, *wobei* $A(B'/B)$ *eine beliebige Formel sei, die aus* A *durch Ersetzung beliebig vieler Vorkommen von* B *durch* B' *entsteht.*

Beweis. Induktion über den Aufbau von A. Für Tempus-freie Formeln übernehmen wir den Beweis des Satzes für **KL**. Der Induktionsschritt ist jetzt zu erweitern durch die Betrachtung von Formeln der Gestalt $A = \mathsf{H}D$ und $A = \mathsf{G}D$. Die Fälle sind gleich und wir betrachten nur den ersten.

Fall $A = \mathsf{H}D$. Wir nehmen an, $\vdash B \leftrightarrow B'$ und B kommt in D vor. Dann gilt die Induktionsannahme $\vdash D \leftrightarrow D(B/B')$. Daraus folgt unmittelbar aufgrund von Satz 5, daß $\vdash \mathsf{H}D \leftrightarrow \mathsf{H}D(B/B')$. ■

Schließlich sind zwei Dualitätsphänomene zu beobachten. Die Dualität von $\vee$ und $\wedge$ drückt sich in deren einfacher Interdefinierbarkeit mittels der Negation aus (DeMorgan-Gesetze!). Da man sich im endlichen Fall die Quantoren $\exists$ und $\forall$ als große Disjunktionen bzw. Konjunktionen denken kann, verhalten diese Quantoren sich in ähnlichem Sinne dual zueinander wie die beiden Junktoren. Wenn wir an die Interpretation von F und G in Zeitmodellen denken, dann erkennen wir in diesen Operatoren unschwer einen Existenz- und einen Allquantor; Gleiches gilt für P und H. Wir dürfen also erwarten, daß diese Paare von Zeitoperatoren ebenfalls dual zueinander stehen.

- Wenn wir in einer Formel A alle Atome gegen ihre Negationen, und die Operatoren $\wedge$ und $\vee$, G und F, und H und P gegenseitig vertauschen, dann nennen wir die so enstandende Formel A^d *dual* zu A. Offensichtlich gilt $A^{dd} = A$.[4]

Eine einfache Induktion über A zeigt, daß

$$A^d \leftrightarrow \neg A. \tag{D}$$

Beweis. Die Basis $A = P$ ist durch $P^d = \neg P$ gegeben. Wir nehmen nun an, (D) gelte für alle Formeln kürzer als A (IA) und betrachten den Fall $A = \neg B$. Da $B^d \leftrightarrow \neg B$ nach (IA), so haben wir $(\neg B)^d \leftrightarrow B^{dd}$, d.h. $(\neg B)^d \leftrightarrow B \leftrightarrow \neg(\neg B)$ wie gewünscht. Im Fall $A = B \vee C$ haben wir $(B \vee C)^d = B^d \wedge C^d \leftrightarrow \neg B \wedge \neg C$ nach Def. und (IA). Also ist $(B \vee C)^d \leftrightarrow \neg(B \vee C)$. Schließlich noch der Fall $A = \mathsf{P}B$: $(\mathsf{P}B)^d = \mathsf{H}B^d \leftrightarrow \mathsf{H}\neg B \leftrightarrow \neg\mathsf{P}B$. ∎

Da nun $A \to B$ und $\neg B \to \neg A$ äquivalente Formeln sind, so folgt aus (D) unmittelbar:

SATZ 7. (Dualität) $\vdash A \to B$ *gdw* $\vdash B^d \to A^d$ *in allen* $\mathbf{K}_t^+$.

Wir haben oben den Begriff der Erweiterung einer Logik so eingeführt, daß alle Erweiterungen der klassischen Logik diese Eigenschaft haben: Wenn eine Formel ein Theorem der Logik ist, dann ist es auch jede Formel der gleichen Form. Dahinter steht der Gedanke, daß logische Wahrheit "formal" ist. Es kann nicht sein, daß einem Satz die Eigenschaft logischer Wahrheit zugesprochen und einem Satz der gleichen Form abgesprochen wird. In der Zeitlogik haben wir es mit der Form der Zeit zu tun. Unter der Annahme, daß die zukünftige Zeit in der Form gleich verläuft wie die vergangene Zeit, müssen wir erwarten, daß jeder in die Zukunft gerichteten zeitlogischen Wahrheit eine strukturell gleiche logische Wahrheit entspricht, die sich auf die Vergangenheit richtet, und umgekehrt.

[4] Zuvor lösen wir in A andere Junktoren nach ihren Definitionen auf so, daß z.B. $A \to B$ zu $\neg A \vee B$ wird. Im Kapitel III über Modallogik geben wir eine induktive Definition der Dualfunktion für alle gebräuchlichen Junktoren.

- Wenn in einer Formel A alle Vorkommen von G durch H (und alle Vorkommen von F durch P) ersetzt werden, dann *spiegelt* die so entstandene Formel A^s die Formel A und *vice versa*; wenn in einer Formel A keine Tempora vorkommen, dann sei $A^s = A$. Wir nennen eine Zeitlogik *spiegelnd*, wenn auch das Spiegelbild jedes Theorems ein Theorem ist.

Unter der erwähnten Annahme wünschen wir uns also, daß die Logik $\mathbf{K}_t$ und jede ihrer zeitlogischen Erweiterungen spiegelnd ist. Der nächste Satz zeigt, daß mit $\mathbf{K}_t$ ein guter Anfang gemacht wurde.

SATZ 8. (Spiegelung) $\mathbf{K}_t$ *ist spiegelnd, d.h.* $\vdash A$ *gdw* $\vdash A^s$.

BEWEIS. Induktion über die Länge einer Ableitung von A. Die Spiegelung jedes Axioms ist ebenfalls ein Axiom. Wenn in einer Ableitung aufgrund einer der beiden zeitlogischen Regeln auf eine Konklusion B geschlossen wurde, dann kann aufgrund der jeweils anderen Regel ebenso auf B^s geschlossen werden. Wenn aufgrund von MP aus $B \to C$ und B auf C geschlossen wurde, dann gibt es nach der Induktionsannahme ableitbare Formeln $(B \to C)^s$ $[= B^s \to C^s]$ und B^s so, daß ebenfalls per MP auf C^s geschlossen werden darf. ∎

Die Forderung, daß eine Zeitlogik spiegelnd sein soll, ist nicht in demselben Maße zwingend wie die Forderung, daß sie – wie jede Logik – formal sein soll. So wird manchmal gesagt, die Zukunft sei, im Gegensatz zur Vergangenheit, "offen". Was immer damit im Einzelnen gemeint sein mag, eine Möglichkeit, diese Vorstellung zu modellieren besteht darin, daß es für jeden Punkt a zwar eine linear angeordnete Vergangenheit gibt – die "Geschichte" von a –, am Punkt a sich die Zukunft jedoch verzweigt in verschiedene Möglichkeiten, wie es von a aus weitergehen könnte. Das ist offensichtlich ein asymmetrisches Bild der Zeit, für das die Forderung der Spiegelbarkeit zeitlogischer Wahrheit unbillig wäre. Auf verzweigte Zeitstrukturen und ihre Logik werden wir im Abschnitt 12 zurückkommen.

Übersetzung in die Prädikatenlogik. Die Wahrheitsbedingung für den Vergangenheitsoperator,

$$(\mathsf{H}) \qquad a \models \mathsf{H}A \text{ gdw } \forall b : b \prec a \Rightarrow b \models A$$

deutet auf der rechten Seite eine Formel an, die wir in einer prädikatenlogischen Sprache so ausdrücken können:

$$\forall b(b \prec a \to Ab)$$

Ähnliches gilt für die anderen Zeitoperatoren. Ganz allgemein, können wir eine zeitlogische Sprache $\mathcal{L}_t$ mit **H** und **G** (und daraus definierbaren Operatoren) in eine prädikatenlogische (PL) Sprache $\mathcal{L}_1$ der ersten Stufe übersetzen. Dazu benötigen wir in $\mathcal{L}_1$ eine zweistellige Relation $<$ sowie genügend viele Prädikate so, daß sich die Atome in $\mathcal{L}_t$ injektiv auf die Prädikate in $\mathcal{L}_1$ abbilden lassen. Der Einfachheit halber nehmen wir an, daß es zwischen den Atomen und den Prädikaten eine Bijektion gibt; es sei $(\underline{\ \ })$ eine solche Bijektion. Wenn P ein Atom ist, dann übersetzen wir: $\underline{P}x_0$. D.h. wir übersetzen die Aussage P in eine Eigenschaft $\underline{P}$ von Zeitpunkten, nämlich die Eigenschaft, daß P zu diesem Zeitpunkt x_0 wahr ist. Die Übersetzung der Atome wird dann wie folgt erweitert zu einer *Standardübersetzung* $(\ )^*$ aller zeitlogischen Formeln in die Sprache der PL erster Stufe:

$$\begin{aligned}
P^* &= \underline{P}x_0 \\
(\neg A)^* &= \neg A^* x_0 \\
(A \wedge B)^* &= A^* x_0 \wedge B^* x_0 \\
(\mathsf{H}A)^* &= \forall x(x < x_0 \rightarrow A^*[x/x_0]) \\
(\mathsf{G}A)^* &= \forall x(x_0 < x \rightarrow A^*[x/x_0])
\end{aligned}$$

In den letzten beiden Bedingungen muß mit x eine "frische" Variable, d.h. eine, die in A nicht vorkommt, gewählt werden.

Es dürfte unmittelbar einleuchten, daß eine solche Übersetzung den Sinn der zeitlogischen Formeln in der Prädikatenlogik gut wiedergibt. Wir können diese Einsicht wie folgt präzisieren. Es sei $\mathcal{M}_t = (T, \prec, I)$ ein Kripke-Modell für die zu übersetzende Sprache $\mathcal{L}_t$. Nun betrachte man eine Interpretation I' in $(T, \prec)$ der PL-Sprache $\mathcal{L}_1$ in welche übersetzt werden soll; dabei sei $I'(<) = \prec$, und $I'(\underline{P}) = I(P)$. (Da wir annehmen, daß $(\underline{\ \ })$ bijektiv ist, ist diese Art der Bezugnahme auf Prädikate in $\mathcal{L}_1$ in Ordnung.) Schließlich sei h eine Belegung der Variablen in $(T, \prec, I')$. Dann gilt folgendes:

$$a \models A \text{ in } (T, \prec, I) \text{ gdw } h \models A^* \text{ in } (T, \prec, I'), \forall h \text{ mit } h(x_0) = a,$$

und ferner

$$\models A \text{ in } (T, \prec, I) \text{ gdw } \models A^* \text{ in } (T, \prec, I').$$

Die Sätze $(1')$ und $(2')$ auf p. 31 sind unschwer als Resultate einer Standardübersetzung zu erkennen. Satz $(1')$ kommt mit zwei Variablen, x_0 und x, aus. In Satz $(2')$ haben wir drei Variablen verwendet. Wir betrachten nur das erste Disjunkt:

$$(2') \qquad \exists x(x_0 < x \wedge \underline{P}x \wedge \exists y(x < y \wedge \underline{Q}y)).$$

Die Verwendung der dritten Variablen y ist jedoch unnötig. Die freie Variable x_0 soll hier mit dem (wechselnden) Evaluierungszeitpunkt belegt werden können. Dazu muß x_0 nur an der in (2′) gezeigten Stelle der Formel frei bleiben; an anderer Stelle können wir dieselbe Variable anders verwenden, d.h. z.B. so binden:

$$(2'') \qquad \exists x(x_0 < x \wedge \underline{P}x \wedge \exists x_0(x < x_0 \wedge \underline{Q}x_0)).$$

Tatsächlich läßt sich die Anzahl der Variablen in der Übersetzung einer zeitlogischen Formel durch geschicktes Recycling (wie im Beipiel illustriert) immer auf zwei reduzieren. (Das zeigt eine einfache Induktion über den Aufbau einer zeitlogischen Formel und ihre jeweilige Übersetzung.) Es sei $\mathbf{T}$ eine Zeitlogik und T^* sei die Übersetzung der zeitlogischen Postulate von $\mathbf{T}$ in eine Sprache erster Stufe mit höchstens zwei Variablen. $\mathbf{PL}_1^2 + T^*$ sei das Zwei-Variablen-Fragments $\mathbf{PL}_1^2$ der Prädikatenlogik erster Stufe, erweitert um T^*. Dann ist eine Formel A genau dann ein Theorem von $\mathbf{T}$, wenn A^* ein Theorem von $\mathbf{PL}_1^2 + T^*$ ist. Da jede Zwei-Variablen-Erweiterung von $\mathbf{PL}_1^2$ entscheidbar ist, so ist auch $\mathbf{T}$ entscheidbar.

8. Richtige Zeit: Erweiterungen von $\mathbf{K}_t$

Nicht *jede* durch eine zweistellige Relation $\prec$ angeordnete Menge von Punkten kann als eine Repäsentation von Zeit dienen. Wir erwarten, daß die Menge mit der Relation, d.h. die Struktur $(T, \prec)$, bestimmte Eigenschaften erfüllt. Nur in Strukturklassen mit den richtigen Eigenschaften wollen wir unserer formalen Zeitsprache interpretieren.

Je stärker die Bedingungen sind, die wir von den zu betrachtenden Zeitstrukturen fordern, desto kleiner wird die Modellklasse über die wir in der Definition einer gültigen Formel quantifizieren, und *ergo* desto größer wird die Chance einer Formel gültig zu sein. Um die Menge der logischen Wahrheiten in Strukturen mit zusätzlichen Bedingungen axiomatisch einzufangen, müssen wir also die Axiomatik von $\mathbf{K}_t$ um zusätzliche Schemata erweitern. In diesem Abschnitt wollen wir untersuchen, wie systematisch der Zusammenhang von Strukturbedingungen einerseits und Axiomenschemata andererseits ist. (Wir könnten auch weitere Regelschemata betrachten, werden aber hier dem Usus folgen, Erweiterungen von $\mathbf{K}_t$ allein mit Hilfe von Formelschemata darzustellen.)

- Unter einer *Strukturbedingung* wollen wir hier eine Eigenschaft der Relation $\prec$ in Kripke-Strukturen verstehen, die sich durch Quantifikation über Zeitpunkte ausdrücken läßt (Eigenschaft erster Stufe). Meist lassen wir die Quantifikation implizit. D.h. statt "$\forall ab(a \prec b \Rightarrow b \not\prec$

a)" schreiben wir einfacher "$a \prec b \Rightarrow b \not\prec$" und denken uns die Allgemeinheit hinzu.

- Wenn (Kond) eine Strukturbedingung ist, dann drücken wir durch $Kond(\mathcal{K})$ aus, daß die Kripke-Struktur $\mathcal{K}$ die Bedingung (Kond) erfüllt; $Kond(\mathbb{K})$ ist die Menge der Kripke-Strukturen, für die (Kond) gilt.

Definition 9. Axiom A *korrespondiert* mit Bedingung (Kond) gdw für alle Strukturen $\mathcal{K} \in \mathbb{K}$ gilt: $\mathcal{K} \models A$ gdw $Kond(\mathcal{K})$. Wenn A und (Kond) korrespondieren, dann *definiert* A die Strukturklasse $Kond(\mathbb{K}) = \{\mathcal{K} \in \mathbb{K} : Kond(\mathcal{K})\}$ in $\mathbb{K}$.

Frage 1: Gibt es für jede Strukturbedingung (Kond) ein korrespondierendes Zeitaxiom A? D.h., (a) wenn $Kond(\mathcal{K})$, dann $\mathcal{K} \models A$ und (b) wenn $\mathcal{K} \models A$, dann $Kond(\mathcal{K})$?

Antwort: Sicher nicht! Einige Bedingungen sind nicht durch Formeln definierbar. Das liegt daran, daß die Sprache erster Stufe wesentlich ausdrucksstärker ist als die Zeitsprache – die Standardübersetzung funktioniert nur in eine Richtung! Zum Beipiel könnten wir Strukturen unter die Bedingung stellen, eine bestimmte Anzahl von Punkten zu enthalten. Eine solche Bedingung läßt sich prädikatenlogisch leicht ausdrücken; in unserer Sprache mit **H** und **G** ist diese Bedingung jedoch nicht definierbar. Man betrachte die Bedingung (Eins), daß T eine Einermenge sei, d.h. die Bedingung $\exists x : x \in T \;\&\; \forall y : y \in T \Rightarrow x = y$. In diesem Fall gibt es keine Formel A derart, daß für alle Strukturen $\mathcal{K}$ gilt: Wenn $\mathcal{K} \models A$, dann $Eins(\mathcal{K})$.

Ein interessanteres Beispiel ist die Bedingung

(Lin) Linearität $a \prec b$ oder $a = b$ oder $b \prec a$

Auch diese Bedingung ist, wie wir gleich sehen werden, nicht durch ein Zeitaxiom definierbar.

Frage 2: Gibt es für jedes Zeitaxiom eine korrespondiernde Strukturbedingung? D.h., gegeben ein Formelschema A, gibt es dann eine Bedingung (Kond) so, daß $\mathcal{K} \models A$ gdw $Kond(\mathcal{K})$?

Antwort: Auch nicht. Einige Formeln definieren keine Bedingungen. Das mag angesichts der Standardübersetzung zunächst überraschend sein. Wir werden ein Beispiel für eine solche Formel kennenlernen.

Strukturbedingungen. Hier ist eine Liste der wichtigsten Bedingungen für Zeitstrukturen. Zusammengenommen fordern sie, daß die Zeitordnung transitiv ist, daß sie in beide Richtungen linear verläuft und unendlich ist, und daß sie kontinuierlich (dicht) ist.

(Trans)	Transitivität	$a \prec b \;\&\; b \prec c \;\Rightarrow\; a \prec c$
(R-Lin)	R-Linearität	$a \prec b \;\&\; a \prec c \;\Rightarrow$ $b \prec c$ oder $b = c$ oder $c \prec b$
(L-Lin)	L-Linearität	$a \prec c \;\&\; b \prec c \;\Rightarrow$ $a \prec b$ oder $a = b$ oder $b \prec a$
(R-Erw)	R-Erweiterbarkeit	$\exists b : a \prec b$
(L-Erw)	L-Erweiterbarkeit	$\exists a : a \prec b$
(Dicht)	Dichte	$a \prec c \;\Rightarrow\; \exists b : a \prec b \;\&\; b \prec c$

In der Liste fehlt die in der Antwort auf die erste Frage erwähnte stärkere Formulierung der Bedingung der Linearität,

(Lin)	Linearität	$a \prec b$ oder $a = b$ oder $b \prec a$

(Lin) impliziert natürlich (R-Lin) und (L-Lin), jedoch nicht umgekehrt, wie die folgende Struktur $\mathcal{P}$ (für "parallel") zweier unverbundener Ketten (die auch die anderen Bedingungen erfüllt) zeigt:

$$\mathcal{P}: \begin{array}{c} \cdots\ a\bullet \longrightarrow b\bullet \longrightarrow c\bullet\ \cdots \\ \cdots\ a'\bullet \longrightarrow b'\bullet \longrightarrow c'\bullet\ \cdots \end{array}$$

Die Parallelstruktur $\mathcal{P}$ erfüllt L- und R-Linearität aber nicht (Lin). Die Teilstrukturen

$$\mathcal{P}_1 : \ \cdots\ a\bullet \longrightarrow b\bullet \longrightarrow c\bullet\ \cdots \qquad \mathcal{P}_2 : \ \cdots\ a'\bullet \longrightarrow b'\bullet \longrightarrow c'\bullet\ \cdots$$

sind natürlich beide linear. Nun ist es aber so, daß in $\mathcal{P}$ und in $\mathcal{P}_1$ (wie auch in $\mathcal{P}_2$) dieselben Formeln gültig sind (siehe z.B. [27, pp. 146f.]). Also gibt es keine zeitlogische Formel, die zwischen der LR-linearen Struktur $\mathcal{P}$ und der linearen Struktur $\mathcal{P}_1$ unterscheiden kann; d.h. es gibt keine Formel A derart, daß wenn $\mathcal{K} \models A$, dann $Lin(\mathcal{K})$. Die Bedingung der Linearität ist nicht durch eine Formel definierbar.

Korrespondenzen. An einem einfachen Beispiel wollen wir nun zeigen, wie man die Korrespondenz zwischen einem Axiom und einer Strukturbedingung nachweist. Dazu gehen wir von einem Axiom aus, das für eine Zeitlogik *keine* gute Wahl ist,

(T**H**) $\qquad \mathsf{H}A \to A,$

was früher immer wahr war, ist wahr. Per (Spiegelung) ist T**H** übrigens äquivalent zu T**G** (Satz 8), und beide Schemata definieren dieselbe Strukturbedingung, wie wir gleich sehen werden.

Satz 10. *Das Axiom* $\mathsf{H}A \to A$ *korrespondiert mit der Strukturbedingung* (Refl) $a \prec a$.

Beweis. Zu zeigen ist: (LR) wenn $\mathcal{K} \models \mathsf{H}A \to A$, dann *Refl*($\mathcal{K}$); und (RL) umgekehrt.

LR: Angenommen eine Struktur $(T, \prec)$ sei nicht reflexiv, d.h. T enthalte einen Punkt a mit (1) $a \not\prec a$. Auf einer solchen Struktur wählen wir ein Modell $(T, \prec, I)$ so, daß für mindestens ein Atom P

$$\forall x : I(P, x) = 1 \text{ gdw } x \prec a. \tag{2}$$

I wird dann zu $\models$ wie üblich (p. 36) fortgesetzt. Wir haben also unmittelbar aus (2),

$$\forall x : x \prec a \;\Rightarrow\; x \models P, \tag{3}$$

und also (4) $a \models \mathsf{H}P$. Aus der Annahme (1) und (2) (nun in umgekehrter Richtung) folgt aber (5) $a \not\models P$. Hier ist eine Skizze des Modells:

$$\underbrace{\cdots \longrightarrow \bullet \longrightarrow \bullet}_{P} \longrightarrow \begin{array}{l} \bullet\, a \\ \mathsf{H}P \\ \neg P \end{array}$$

Aus (4) und (5) folgt $a \not\models \mathsf{H}P \to P$. In einer nicht-reflexiven Struktur kann also T**H** falsifiziert werden.

RL: Wir betrachten eine reflexive Struktur $\mathcal{K}$ und zeigen, daß $a \models \mathsf{H}A \to A$ für einen beliebigen Punkt in einem beliebigen Modell auf $\mathcal{K}$. Wir nehmen also an, daß $a \models \mathsf{H}A$ (und zeigen, daß $a \models A$). Die Annahme bedeutet, daß $\forall x : x \prec a \;\Rightarrow\; x \models A$, woraus aufgrund der Reflexivität von $\prec$ das gewünschte $a \models A$ folgt. ■

Tansitivität

Im weiteren betrachten wir Korrespondenzen, die besser in die Zeitlogik passen als diejenige, welche wir in Satz 10 nachgewiesen haben. Wir beginnen mit der Bedingung der *Transitivität* der Relation $\prec$ und zeigen, daß diese durch das Axiom 4**G** definiert ist.

(Trans)	Transitivität	$a \prec b \;\&\; b \prec c \Rightarrow a \prec c$
4**G**.		$\mathsf{G}A \to \mathsf{GG}A$

Der Nachweis, daß 4**G** in *Trans*($\mathbb{K}$) gültig ist, sei dem Leser überlassen (nach dem Muster des RL-Teils des obigen Beweises).

Daß die Gültigkeit von 4G die Eigenschaft der Transitivität erzwingt, zeigen wir wie folgt. Wir betrachten ein Modell auf einer Struktur, in der $\prec$ nicht transitiv ist (und zeigen, daß in diesem Modell eine Instanz von 4G falsifiziert werden kann). Das Modell bestehe aus drei Punkten mit $b \models P$ und $c \models \neg P$:

$$\begin{array}{ccccc} a\bullet & \longrightarrow & b\bullet & \longrightarrow & c\bullet \\ & & P & & \neg P \end{array}$$

Wir nehmen ferner an, daß $a \not\prec c$! In diesem Modell gilt $a \models \mathsf{G}P$, jedoch nicht $b \models \mathsf{G}P$ und also auch nicht $a \models \mathsf{GG}P$; somit $a \not\models \mathsf{G}P \rightarrow \mathsf{GG}P$.

Übrigens würde statt 4G auch irgendeines der Schemata

4H	$\mathsf{H}A \rightarrow \mathsf{HH}A$
4P	$\mathsf{PP}A \rightarrow \mathsf{P}A$
4F	$\mathsf{FF}A \rightarrow \mathsf{F}A$

dem Zweck dienen. Daß alle vier Versionen des 4-Schemas äquivalent sind, können wir durch Modellbetrachtungen bestätigen. Alternativ betrachten wir z.B. das System $\mathbf{K}_t + 4\mathsf{G}$ und erinnern uns, daß in diesem System auch das duale Schema 4F ableitbar sein muß (Satz 7). Sodann leiten wir in dem System 4P ab. [5]

(1)	$\mathsf{GP}A \rightarrow \mathsf{GGP}A$	4P
(2)	$A \rightarrow \mathsf{GP}A$	B
(3)	$A \rightarrow \mathsf{GGP}A$	1,2 **KL**
(4)	$\mathsf{PP}A \rightarrow \mathsf{PPGGP}A$	3, Monotonie
(5)	$\neg\mathsf{GP}A \rightarrow \mathsf{HF}\neg\mathsf{GP}A$	B
(6)	$\neg\mathsf{HF}\neg\mathsf{GP}A \rightarrow \mathsf{GP}A$	5, **KL**
(7)	$\mathsf{PGGP}A \rightarrow \mathsf{GP}A$	6, Def.
(8)	$\mathsf{PPGGP}A \rightarrow \mathsf{PGP}A$	7, Monotonie
(9)	$\mathsf{PP}A \rightarrow \mathsf{PGP}A$	4, 8 , **KL**
(10)	$\mathsf{H}A \rightarrow \mathsf{HFH}A$	B
(11)	$\mathsf{PGP}A \rightarrow \mathsf{P}A$	10, Dualität
(12)	$\mathsf{PP}A \rightarrow \mathsf{P}A$	9,11, **KL**

Schließlich stellen wir fest, daß aufgrund der Dualität von 4P und 4H auch dieses letzte Schema ableitbar sein muß. Genauso können wir mit jeder Erweiterung von $\mathbf{K}_t$ um eines der vier Schemata verfahren; so stellen sich die Schemata in $\mathbf{K}_t$ als äquivalent heraus.

[5] Die Ableitung folgt Burgess in [43, p. 28].

Linearität

Die Bedingung, daß die Zeit linear verläuft, besteht aus zwei Teilbedingungen: Erstens, daß die Vergangenheit eine Kette von Zeitpunkten bildet, und zweitens, daß die Zukunft eine solche Kette bildet (und beide Teile derselben Kette sind). Wenn wir z.B. geschichtliche Zeit betrachten, dann nehmen wir meist an, daß die erste aber nicht die zweite Bedingung zu trifft. Die Vergangenheit ist abgeschlossen; sie ist eine Verkettung von Ereignissen. Die Zukunft ist jedoch offen; in ihr verzweigen sich die Ereignispfade in verschiedene Möglichkeiten, wie es weiter gehen könnte.

Hier ist die Bedingung der Linearität der Zukunft zusammen mit dem korrespondierenden Formelschema:

(R-Lin) R-Linearität $a \prec b \mathbin{\&} a \prec c \;\Rightarrow\; b \prec c$ oder $b = c$ oder $c \prec b$

H_r. $\mathsf{F}A \wedge \mathsf{F}B \to \mathsf{F}(A \wedge \mathsf{F}B) \vee \mathsf{F}(A \wedge B) \vee \mathsf{F}(\mathsf{F}A \wedge B)$

$\mathrm{H}_r \Rightarrow$ (R-Lin). Dieses Gegenmodell zu (R-Lin) hatten wir schon oben betrachtet:

$$\begin{array}{ccc} & & \bullet b \quad A, \neg B \\ & \nearrow & \\ a\bullet & & \\ & \searrow & \\ & & \bullet c \quad B, \neg A \end{array}$$

Es ist leicht zu sehen (und hier mit Buchstaben angedeutet), daß man in diesem Modell Einsetzungen für A und B wählen kann so, daß das Antezedens von $\mathrm{H_r}$ wahr und das Konsequens falsch ist. Also erzwingt die Wahrheit des Schemas in einem Modell, daß das Modell rechts-linear ist.

(R-Lin) $\Rightarrow \mathrm{H}_r$. Umgekehrt ist die Bedingung (R-Lin) hinreichend, um in einem Modell das Schema H_r zu verifizieren. Dazu nehmen wir an, daß (1) $a \models \mathsf{F}A$ und $a \models \mathsf{F}B$ (und zeigen unter Verwendung der Bedingung, daß $a \models$ *Konsequent*). Dann (2) $\exists b : a \prec b$ mit $b \models A$ und (3) $\exists c : a \prec c$ mit $c \models B$. Nach (R-Lin) haben wir jetzt (4a) $b \prec c$ oder (4b) $b = c$ oder (4c) $c \prec b$. Unter den Annahmen genügt jeweils eines dieser Disjunkte, um zu zeigen, daß (a) $a \models \mathsf{F}(A \wedge \mathsf{F}B)$ oder (b) $a \models \mathsf{F}(A \wedge B)$ oder (c) $a \models \mathsf{F}(\mathsf{F}A \wedge B)$.

Die Bedingung der Linearität der Vergangenheit und das dazu passende Formelschema verhalten sich spiegelbildlich zu dem Paar, das wir gerade betrachtet haben:

(L-Lin) L-Linearität $a \prec c \mathbin{\&} b \prec c \Rightarrow a \prec b$ oder $a = b$ oder $b \prec a$

H_ℓ. $\mathsf{P}A \wedge \mathsf{P}P \to \mathsf{P}(A \wedge \mathsf{P}B) \vee \mathsf{P}(A \wedge B) \vee \mathsf{P}(\mathsf{P}A \wedge B)$

Der Nachweis der Korrespondenz ist völlig analog zu dem gerade besprochenen Paar.

Aufgrund der Dualitätsbeobachtung in Satz 7 können wir H_r und H_ℓ auch jeweils durch ihre dualen Formen ersetzen:

H_r^d $\quad \mathsf{G}(A \vee \mathsf{G}B) \wedge \mathsf{G}(A \vee B) \wedge \mathsf{G}(\mathsf{G}A \vee B) \to \mathsf{G}A \vee \mathsf{G}B$

H_ℓ^d $\quad \mathsf{H}(A \vee \mathsf{H}B) \wedge \mathsf{H}(A \vee B) \wedge \mathsf{H}(\mathsf{H}A \vee B) \to \mathsf{H}A \vee \mathsf{H}B$

Erweiterbarkeit (Unendlichkeit)

Die Zeit ist unendlich: Sie hat keinen Anfang und kein Ende. Auch dies ist eine Konjuktion zweier Behauptungen, nämlich , daß die Zeit sich unendlich fortsetzt, erstens, in die Vergangenheit (L-Erw) und, zweitens, in die Zukunft (R-Erw). Die zu diesen Bedingungen korrespondierenden Formelschemata sind

(R-Erw)	R-Erweiterbarkeit	$\exists b : a \prec b$
D_r.		$\mathsf{G}A \to \mathsf{F}A$
(L-Erw)	L-Erweiterbarkeit	$\exists a : a \prec b$
D_ℓ.		$\mathsf{H}A \to \mathsf{P}A$

Das kleinste Gegenmodell zu D_r besteht aus nur einem Punkt:

$$\begin{array}{c} a\bullet \\ \mathsf{G}A, \neg\mathsf{F}A \end{array}$$

Man beachte, daß $a \models \mathsf{G}A$ gdw $\forall b : a \prec b \Rightarrow b \models A$. Da es kein b mit $a \prec b$ gibt, ist die Bedingung für $a \models \mathsf{G}A$ auf triviale Weise erfüllt während die Bedingung für $a \models \mathsf{F}A$ nicht erfüllt ist. Wenn also D_r in einem Modell wahr sein soll, dann muß die Bedingung (R-Erw) gelten.

Für die umgekehrte Richtung nehmen wir an, daß $\mathsf{G}A$ an einem Punkt a wahr ist, d.h. wenn a eine Zukunft hat, dann ist A wahr an allen Punkten dieser Zukunft. Nun garantiert (R-Lin) genau diese Bedingung, d.h. a hat eine Zukunft. Also gibt es Punkte in der Zukunft von a, an welchen A wahr ist.

Mit dem Nachweis der Korrespondenz zwischen (L-Erw) und D_r verfahren wir analog.

Dichte

Die Zeit ist nicht nur in die Vergangenheit und Zukunft unendlich (erweiterbar), sondern auch zwischen je zwei Zeitpunkten. Wir stellen uns die

Zeit als ein *Kontinuum* vor, d.h. so, daß zwischen je zwei Zeitpunkten immer noch ein weiterer Zeitpunkt liegt. Die Zeitpunkte sind dicht geordnet, wie die reellen Zahlen **R**.

(Dicht)	Dichte	$a \prec c \Rightarrow \exists b : a \prec b \;\&\; b \prec c$
4c**G**.		$\mathsf{GG}A \to \mathsf{G}A$

Die Formel 4c**G** ist die Umkehrung von 4**G**, welche die Eigenschaft der Transitivität ausdrückt. Aufgrund von Überlegungen wie wir sie bei der Besprechung von 4**G** angestellt haben, ist 4c**G** äquivalent zu jedem der folgenden Schemata:

4c**H**	$\mathsf{HH}A \to \mathsf{H}A$
4c**P**	$\mathsf{P}A \to \mathsf{PP}A$
4c**F**	$\mathsf{F}A \to \mathsf{FF}A$

Um uns davon zu überzeugen, daß das Schema die Bedingung fordert, betrachten wir ein Modell mit nur drei Punkten:

$$\begin{array}{ccccc} a\bullet & \longrightarrow & b\bullet & \longrightarrow & c\bullet \\ \mathsf{GG}P & & \mathsf{G}P & & \\ \neg\mathsf{G}P & & \neg P & & P \end{array}$$

Ein Atom P sei so interpretiert, daß $b \not\models P$ und $c \models P$. Dann $b \models \mathsf{G}P$ und also $a \models \mathsf{GG}P$. Da jedoch die Zukunft von a nur aus b besteht, so haben wir $a \not\models \mathsf{G}P$.

Für die umgekehrte Korrespondenzrichtung (die Bedingung ist hinreichend für das Schema) nehmen wir an, daß (1) $a \models \mathsf{GG}A$ (und zeigen, daß $a \models \mathsf{G}A$). Wir nehmen weiter an, daß (2) $a \prec b$ (beliebiges b) (und zeigen, daß $b \models A$). Packen wir die Wahrheitsbedingung von (1) aus, dann haben wir (3): $\forall xy : a \prec x$ und $x \prec y \Rightarrow y \models A$. Aber aus (2) erhalten wir per (Dicht): (4) $\exists x : a \prec x$ und $x \prec b$. (4) gibt uns eine Instanz des Antezedens' von (3); daher $b \models A$, wie gewünscht.

9. Definierbarkeit und Tempora in $\mathbf{K}_t^*$

$\mathbf{K}_t^* = \mathbf{K}_t + \text{HD4!}$ sei die Erweiterung von $\mathbf{K}_t$ mit den folgenden Axiomen:

H_r	$\mathsf{F}A \wedge \mathsf{F}B \to \mathsf{F}(A \wedge \mathsf{F}B) \vee \mathsf{F}(A \wedge B) \vee \mathsf{F}(\mathsf{F}A \wedge B)$
H_ℓ	$\mathsf{P}A \wedge \mathsf{P}B \to \mathsf{P}(A \wedge \mathsf{P}B) \vee \mathsf{P}(A \wedge B) \vee \mathsf{P}(\mathsf{P}A \wedge B)$
D_r	$\mathsf{G}A \to \mathsf{F}A$
D_ℓ	$\mathsf{H}A \to \mathsf{P}A$
4!	$\mathsf{GG}A \leftrightarrow \mathsf{G}A$

Es sei $\mathbb{S}^*$ die Klasse aller Strukturen, welche die o.g. Eigenschaften erfüllen; also aller transitiven, L- und R-linearen, L- und R-erweiterbaren, sowie dichten Rahmen. Aufgrund der vorhergehenden Überlegungen sind Theoreme dieser Logik genau solche Formeln, die gültig sind in der Strukturklasse $\mathbb{S}^*$, d.h, es gilt der folgende Satz.

Satz 11. $\mathbf{K}_t^* = Th(\mathbb{S}^*)$.

Insofern also $\mathbb{S}^*$ die Struktur der Zeit richtig und vollständig angibt (eine unendlich in die Vergangenheit und Zukunft reichende, dichte Kette von Momenten), beschreibt $\mathbf{K}_t^*$ die Logik der Zeit in unserer Zeitsprache mit den Operatoren H und G.

Definierbarkeit. $\mathbf{K}_t^*$ kommt der Logik der Zeit – so wie sie etwa in der klassischen Physik verstanden wird – ziemlich "nahe"; aber es gibt einige Probleme. Kurz gesagt, fängt $\mathbf{K}_t^*$ auch Strukturen ein, von denen wir nicht sagen wollen, daß sie die **Zeit** repräsentieren. D.h. die Menge $\mathbb{S}^*$ enthält neben den Strukturen, die wir im Sinne haben, wenn wir an die Zeit denken, auch solche, an die wir dabei nicht denken.

Nicht-intendierte Strukturen

- Strukturen mit Parallelzeiten (wie oben beschrieben) sind in $\mathbb{S}^*$ enthalten.
- $\mathbb{S}^*$ enthält neben der intendierten $(\mathbf{R}, \prec)$-Struktur auch die nicht-intendierte $(\mathbf{Q}, \prec)$-Struktur ($\mathbf{R}$: die reellen Zahlen; $\mathbf{Q}$: die rationalen Zahlen). Denn auch $\mathbf{Q}$ erfüllt die Bedingung der Dichte: Zwischen je zwei rationalen Zahlen a und b liegt das arithmetische Mittel $\frac{m+n}{2}$.)
- Auch alle reflexiven Einpunktstrukturen, $(\{a\}, \prec)$ mit $a \prec a$, sind in $\mathbb{S}^*$.
- Ebenso verhält es sich mit den zyklischen Strukturen $(\{a_1, ..., a_n\}, \prec)$ mit $a_1 \prec \cdots \prec a_n \prec a_1$.

Jede dieser nicht-intendierten Strukturklassen $\mathbb{T}$ ist also in der Strukturklasse $\mathbb{S}^*$ enthalten, d.h.

$$\mathbb{T} \subseteq \mathbb{S}^*.$$

Die in $\mathbb{S}^*$ gültigen Formeln sind also auch in $\mathbb{T}$ gültig, d.h.

$$Th(\mathbb{S}^*) \subseteq Th(\mathbb{T}).$$

Da nach Satz 11, $Th(\mathbb{S}^*) = \mathbf{K}_t^*$, so gelten also die Theoreme von $\mathbf{K}_t^*$ für einen weiteren als den intendierten Bereich der **Zeit**. Anders gesagt: $\mathbf{K}_t^*$ gibt zwar logische Eigenschaften von H und G wieder, wenn wir diese als Vergangenheits- und Zukunftsoperatoren in der **Zeit** interpretieren. Aber diese logischen Eigenschaften legen die Interpretation von H und G nicht als Operationen in der **Zeit** fest. Somit drängt sich die Frage auf, ob wir weitere logische Eigenschaften der Operatoren angeben können, d.h. weitere Axiome finden können, welche $\mathbf{K}_t^*$ so erweitern, daß wir genau die Logik der **Zeit** und von keinen davon verschiedenen Strukturen erhalten. Es stellt sich heraus, daß das nicht vollständig gelingen kann.

Irreflexivität

Wir haben oben gesehen, daß es für die Bedingung der Reflexivität aller Punkte ein korrespondierendes Axiom gibt, nämlich $\mathsf{H}A \to A$. Jetzt fragen wir, ob es ein Axiom geben kann, welches Strukturen mit reflexiven Punkten ausschließt. Damit ist folgendes gemeint (vgl. Def. 9): Gibt es ein Axiom X so, daß für alle Kripke-Strukturen $\mathcal{S}$, $\mathcal{S} \models X$ genau dann, wenn $\mathcal{S}$ irreflexiv ist (d.h. für kein $a \in T$ haben wir $a \prec a$)? Wenn das so ist, dann haben wir $\mathbf{K}_t^* + X = Th(Irr(\mathbb{S}^*))$. Unter der Annahme, daß die Früher-Später-Relation der **Zeit** irreflexiv ist, wäre $\mathbf{K}_t^* + X$ einer Logik der Zeit daher ähnlicher als $\mathbf{K}_t^*$.

Wir nehmen einmal an, daß es eine solche Formel X gibt, d.h. für alle Kripke-Strukturen $\mathcal{S}$:

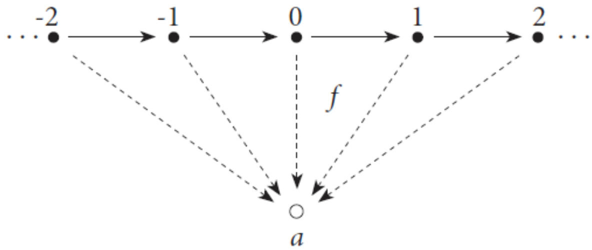
$$(*) \qquad \mathcal{S} \models X \text{ gdw } Irr(\mathcal{S})$$

Die Struktur der ganzen Zahlen, $\mathcal{Z} = (\mathbf{Z}, <)$ ist eine Kripke-Struktur mit der natürlichen irreflexiven Ordnung $<$. Also folgt aus $(*)$,

$$\mathcal{Z} \models X.$$

Nun betrachten wir die Reduzierung f von $\mathcal{Z}$ auf die Einpunkt-Struktur $f(\mathcal{Z}) = (\{a\}, \leq)$ derart, daß

$$(f) \qquad \forall x \in \mathbf{Z} : f(x) = a \text{ und } f(x) \leq f(y) \text{ gdw } x < y.$$



Auch das ist eine Kripke-Struktur. Es ist klar, daß die in (f) definierte Relation $\leq$ reflexiv ist (angedeutet durch den ungefüllten Punkt a). Nun betrachten wir ein Paar von Modellen $\mathcal{M}$ auf $\mathcal{Z}$ und $\mathcal{M}'$ auf $f(\mathcal{Z})$ so, daß für alle $x \in \mathbf{Z}$ und alle Atome P,

$$(P) \qquad x \models_{\mathcal{M}} P \text{ gdw } f(x) \models_{\mathcal{M}'} P.$$

Aus unserer Annahme (1) folgt, daß X an allen Punkten in $\mathcal{M}$ wahr ist. Da der Punkt a in $\mathcal{M}'$ (d.h. $f(x)$ für beliebige $x \in \mathbf{Z}$) *reflexiv* ist, müßte X an diesem Punkt falsch sein. Aber das ist nicht der Fall! Denn durch Induktion über A können wir ganz allgemein zeigen:

Für alle $x \in \mathbf{Z}$, wenn $x \models_{\mathcal{M}} A$, dann $f(x) \models_{\mathcal{M}'} A$.

(Es gilt übrigens auch die Umkehrung.) Entgegen unserer Annahme, kann X also tatsächlich nicht irreflexive von reflexiven Strukturen unterscheiden.

Beweis der Behauptung: Der Fall $A = P$ ist durch (P) gegeben. Die wahrheitsfunktionalen Zusammensetzungen sind Routine. Es sei nun $A = \mathsf{H}B$. Angenommen (1) $x \models \mathsf{H}B$ in $\mathcal{M}$. d.h. $\forall y : x \prec y \Rightarrow y \models B$. Wir nehmen ferner an, daß (2) $f(x) \leq f(y)$ in $\mathcal{M}'$ (und zeigen, daß $f(y) \models B$.) Aus (2) und (f) folgt, daß $x \prec y$ und also aus (1), $y \models B$. Darauf wenden wir die I.H. an und erhalten $f(y) \models B$, wie gewünscht. ■

Die Abbildung f von $\mathcal{Z}$ nach $f(\mathcal{Z})$ ist ein Fall einer *p-morphen Abbildung*. Ohne an dieser Stelle ins Detail zu gehen, gilt für p-Morphismen p von Kripke-Strukturen auf Kripke-Stukturen allgemein (jedoch nicht umgekehrt):

Wenn $\mathcal{S} \models A$, dann $p(\mathcal{S}) \models A$.

Von dieser Beobachtung ausgehend, können wir das Argument oben so abkürzen: Wenn $(*)$ und also $\mathcal{Z} \models X$, dann auch $f(\mathcal{Z}) \models X$. Da aber $\neg Irr(f(\mathcal{Z}))$, so folgt wiederum aus $(*)$, daß $f(\mathcal{Z}) \not\models X$ – Widerspruch. Also kann es keine Formel X geben, welche die Struktureigenschaft der Reflexivität im Sinne von $(*)$ definiert.

Die Modelle $\mathcal{M}$ und $\mathcal{M}'$ sind einander in besonderer Weise ähnlich. Zwischen den Modellen besteht gewissermaßen eine extrinsische und intrinsische Gleichheit der Punkte: Sie unterscheiden sich nicht im Hinblick auf die Übergangsmöglichkeiten zu weiteren Punkten mittels der jeweiligen Relation; und die Atome sind über sie auf gleiche Weise verteilt. Den genauen Sinn dieser zwei Aspekte geben die Bedingungen (f) und (P) an. Diese "Ähnlichkeit" ist eine Erscheinungsform einer allgemeineren Beziehung zwischen Modellen, *Bisimulation* genannt, die wir im Kapitel III über Modallogik eingehender behandeln werden. Die nächste Beobachtung, über Linearität, beruht ebenfalls auf einer Bisimulation zwischen Modellen.

Linearität

Parallelzeiten, d.h. zwei (oder mehr) nebeneinander herlaufende Ketten, können wir durch die Bedingung der Linearität, (Lin), ausschließen:

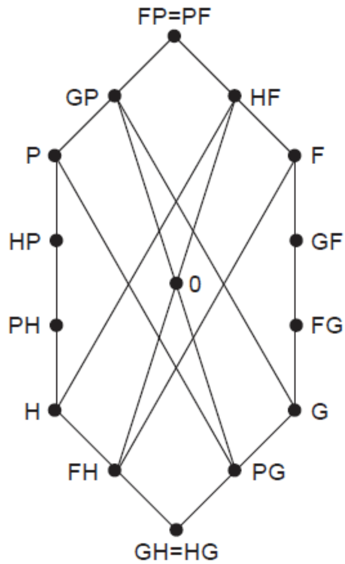
(Lin) $$a \prec b \text{ oder } a = b \text{ oder } b \prec a$$

Aber ist (Lin) durch ein zeitlogisches Axiom definierbar?

Wenn es ein solches Axiom L gäbe, dann müßte es in allen linearen Strukturen gültig sein und in nicht-linearen Strukturen falsch werden können. Wir betrachten jetzt die *disjunkte Vereinigung* $\mathcal{L}_1 \oplus \mathcal{L}_2$ zweier linearer Strukturen $\mathcal{L}_1$ und $\mathcal{L}_2$: Das ist z.B. die parallele aber unverbundene Zusammenfassung zweier Ketten von Momenten zu einer neuen Struktur (vgl. das Bild auf p. 54). Im Gegensatz zu $\mathcal{L}_1$ und $\mathcal{L}_2$ ist die Vereinigung $\mathcal{L}_1 \oplus \mathcal{L}_2$ nicht linear. Wir würden daher erwarten, daß $\mathcal{L}_1 \oplus \mathcal{L}_2 \not\models L$. Aber das ist nicht der Fall. Denn wir können ganz allgemein folgendes beobachten:

Wenn $\mathcal{S}_1 \models A$ *und* $\mathcal{S}_2 \models A$, *dann* $\mathcal{S}_1 \oplus \mathcal{S}_2 \models A$.

Tempora. Daß "es war einmal" und "es wird einmal" verschiedene Zeitmodi (Tempora) audrücken, leuchtet sofort ein. Mit der Beurteilung zusammengesetzter Tempora tuen wir uns jedoch schwer. Selbst in der deutschen Grammatik vorgesehene komplexe Tempora wie das Plusquamperfekt "hatte" und das Futur II "wird einmal gewesen sein", werden in der Umgangssprache eher gemieden. Aber auch für diese Tempora ist klar, daß sie distinkt sind; wir können sie übrigens mit PP bzw FP wiedergeben. Ein nützlicher Aspekt der Zeitlogik und ihrer Semantik besteht darin, daß sie uns erlaubt, beliebig komplexe Tempora sicher zu interpretieren. Ohne eine solche Hilfe können uns in dieser Hinsicht schon relativ einfache Tempora überfordern. Wie sieht es z.B. aus mit "es war immer der Fall, daß A



Der 15-Zeiten-Satz von Hamblin (logische Implikation von unten nach oben).

immer wahr sein wird” einerseits und “es wird immer wahr sein, daß A immer der Fall war” andererseits? Sind das logisch distinkte oder äquivalente Tempora? Und wieviele logisch distinkte Tempora gibt es in $\mathbf{K}_t^*$ überhaupt?

Wie man die Frage angehen kann, verdeutlichen am besten einige beispielhafte Überlegungen.

- Angesichts 4 and 4c sind $\mathsf{G}A$ und $\mathsf{GG}A$ logisch äquivalent. In diesem Sinne macht $\mathbf{K}_t^*$ keinen Unterschied zwischen den zwei Tempora **G** und **GG**.
- Fortgesetzte Iteration per 4 und Reduktion per 4c zeigt, daß

$$\mathsf{G}A \leftrightarrow \mathsf{G}\cdots\mathsf{G}A$$

  für beliebige (endliche) Iterationen von **G**.
- Dasselbe gilt für **F**, **H** und **P**.
- Interessantere Äquivalenzen (siehe [43, p. 30f.]) sind

$$\mathsf{FP}A \leftrightarrow \mathsf{PF}A \text{ und } \mathsf{GH}A \leftrightarrow \mathsf{HG}A.$$

- **FP** (bzw. **PF**) ist das logisch schwächste Tempus. Denn für alle Tempora T (einschl. dem “leeren” oder Gegenwartstempus) gilt: $\vdash TA \to \mathsf{FP}A$. Dagegen ist **GH** das stärkste Tempus: $\mathsf{GH}A \to TA$. ($\mathsf{FP}A$ sagt, daß A irgendwann einmal wahr gewesen sein wird; $\mathsf{GH}A$ sagt, daß A immer durchgehend wahr gewesen sein wird.)

Wieviele logisch distinkte Tempora (in einer Sprache mit $\{\mathsf{P}, \mathsf{H}, \mathsf{F}, \mathsf{G}\}$) verbleiben also in $\mathbf{K}_t^*$ und in welchen Implikationsverhältnissen stehen sie zueinander? Die Frage beantwortet auf erfreulich überschaubare Weise der *Satz von Hamblin* (1958, cf. [232, pp. 45-50] und [43, pp. 29-33]), den wir im Diagramm oben zusammengefaßt haben.

10. Zwei Stellen: Seit und bis

Nach Hamblins Satz können wir in der Logik $\mathbf{K}_t^*$ für die Sprache $\mathbf{L}_t$ 15 logisch verschiedene Tempora ausdrücken. Einige dieser Tempora haben natürliche Entsprechungen in natürlichen Sprachen; darunter, wie schon erwähnt, FP (Futur II) und das PP (Plusquamperfekt). Andere haben keine solchen Entsprechungen. Z.B. das Spiegelbild PF des Futur II: *¿Peter hat gegessen werden?*. Wiederum andere temporale Konstruktionen sind jenseits der Reichweite von $\mathbf{L}_t$. Dazu gehören der Gegenwartsoperator *Jetzt*, den wir im nächsten Abschnitt behandeln werden, und Operationen in denen wir beschreibend auf Zeitpunkte Bezug nehmen, wie in

(1) *Seit* Ulla gegessen hat, ist sie nicht mehr hungrig, oder
(2) *Bis* Ulla gegessen haben wird, wird sie hungrig sein.

Offenbar sind diese beiden Operatoren zweistellig wie ein Konditional. In (1) und (2) drückt der Satz (A) "Ulla ißt" so etwas wie eine temporal geladene Bedingung aus für den Satz (B) "Ulla ist hungrig". Die Verknüpfungen **seit** und **bis** setzen (A) und (B) in ein temporales Verhältnis zueinander. Der Satz (2), also A *bis* B, drückt aus, daß es einen Zeitpunkt in der Zukunft gibt, zu dem Ulla essen wird, und daß sie bis zu diesem Zeitpunkt hungrig sein wird. Ganz allgemein ist A *bis* B jetzt wahr, wenn es einen Zeitpunkt in der Zukunft gibt, zu dem die Bedingung B eintritt, und bis dahin A wahr sein wird. Und A *seit* B ist jetzt wahr, wenn es einen Zeitpunkt in der Vergangenheit gab zu dem B eintrat, und seitdem A der Fall war. In Zeitmodellen können wir das so beschreiben:

(**seit**) $$a \models A \ \mathsf{seit}\ B \text{ gdw } \exists c : c \prec a \text{ und } c \models B, \text{ und } \forall b : c \prec b \prec a \Rightarrow b \models A$$

(**bis**) $$a \models A \ \mathsf{bis}\ B \text{ gdw } \exists c : a \prec c \text{ und } c \models B, \text{ und } \forall b : a \prec b \prec c \Rightarrow b \models A$$

Der aufmerksame Leser wird eine kleine Unstimmigkeit zwischen (1) und der Wahrheitsbedingung (**seit**) bemerkt haben. Der Satz (1) könnte an einem Punkt a nämlich so wahr sein: Ulla hat zum Zeitpunkt a gegessen (B) und ist ab und einschließlich a nicht mehr hungrig (A). Wenn wir aber auf die Klausel (seit) für A seit B schauen, dann sehen wir, daß B an einem Punkt c in der Vergangenheit von a wahr sein muß, und daß die Punkte b, an denen A wahr ist den Punkt a nicht einschließen dürfen. Also etwa: Ulla war satt, **seit** sie gegessen hatte. Offenbar lassen also *bis* und *seit* Varianten zu, je nachdem wie stark Vergangenheits- und Zukunftsaspekte in den Vordergrund gestellt werden. Eine Lesart, die weniger strikt als (seit) ist und für den Satz (1) naheliegender ist, ist diese:

$$(\mathsf{seit}') \qquad a \models A \;\mathsf{seit}'\; B \text{ gdw } \exists c : c \preceq a \text{ und } c \models B, \\ \text{und } \forall b : c \prec b \preceq a \;\Rightarrow\; b \models A,$$

wobei $a \preceq b$ kurz ist für: $a \prec b$ oder $a = b$. Eine ensprechende Modifizierung des strikten **bis** sieht so aus:[6]

$$(\mathsf{bis}') \qquad a \models A \;\mathsf{bis}'\; B \text{ gdw } \exists c : a \preceq c \text{ und } c \models B, \\ \text{und } \forall b : c \preceq b \prec a \;\Rightarrow\; b \models A.$$

Für unsere Zwecke sind die strikten **seit**- und **bis**-Operatoren die bessere Wahl. Denn wir können die nicht-strikten Operatoren seit' und bis' einfach per Definition einführen:

$$A \;\mathsf{seit}'\; B := ((A \;\mathsf{seit}\; B) \wedge A) \vee B$$
$$A \;\mathsf{bis}'\; B := ((A \;\mathsf{bis}\; B) \wedge A) \vee B.$$

Umgekehrt können wir jedoch nicht die strikten aus den nicht-strikten Operatoren zu definieren.

Wenn wir die Klauseln für die zweistelligen Operatoren **seit** und **bis** mit denen für die einstelligen Zeitoperatoren vergleichen, dann fällt auf, daß jene mit *drei* (statt zwei) Variablen über Zeitpunkte formuliert sind. Die Form der beiden Klauseln ist

$$\forall a (F_1 a \leftrightarrow \exists c (F_2 ac \wedge F_3 c \wedge \forall b (F_4 abc \to F_5 b))).$$

Keine der verwendeten Variablen wird für eine Teilformel wieder frei. Also können wir die Anzahl der Variablen nicht durch geschickte Wiederverwendung reduzieren. Im Gegensatz zu einer Sprache in $\{\mathsf{P}, \mathsf{H}, \mathsf{F}, \mathsf{G}\}$ werden für die Standardübersetzung einer zeitlogischen Formeln, in der auch die Operatoren **seit** und **bis** vorkommen, in eine Sprache erster Stufe nicht

[6] Diese Lesart von *bis* wird in Logiken der Programmverifizierung als Until-Operator zugrunde gelegt. Siehe [115] für eine kurze Darstellung solcher Logiken.

weniger als drei Variablen benötigt. Tatsächlich können wir zeigen, daß nicht mehr als drei Variablen benötigt werden; vgl. p. 52. Daher ist eine Formel genau dann ein Theorem einer Zeitlogik mit **seit** und **bis**, wenn deren Standardübersetzung ein Theorem des um die entsprechenden zeitlogischen Axiome erweiterten Drei-Variablenfragments der Prädikatenlogik erster Stufe ist. Im Gegensatz zum Zwei-Variablenfragment ist letzteres jedoch nicht allgemein entscheidbar.

Die neuen Operatoren erweitern nicht nur Ausdruckskraft unserer formalen Sprache; sie können sogar die bisher betrachteten einstelligen Operatoren ersetzen. So besagt $\top$ **bis** A, daß es einen Zeitpunkt in der Zukunft gibt, zu dem A wahr ist und das bis dahin $\top$ wahr ist. Aber $\top$ ist immer wahr. Also ist $\top$ **bis** A genau dann wahr, wenn zukünftig einmal A wahr ist. Das ist genau die Wahrheitsbedingung für $\mathsf{F}A$. Also haben wir

$$\mathsf{F}A \leftrightarrow \top \ \mathsf{bis}\ A.$$

Ebenso stellen wir fest, daß

$$\mathsf{P}A \leftrightarrow \top \ \mathsf{seit}\ A.$$

Die beiden Äquivalenzen können wir im Sinne von Definitionen von **F** und **P** verwenden. So sind die zweistelligen Operatoren ausreichend um alles auszudrücken, was in der Sprache mit den einstelligen Operatoren ausgedrückt werden kann. Da sich aber umgekehrt die zweistelligen nicht aus den einstelligen Operatoren definieren lassen, so ist das Paar (**bis**, **seit**) strikt ausdrucksstärker als das Paar (**P**, **F**) bzw. das funktional äquivalente Paar (**H**,**G**).

11. Zwei Punkte: Jetzt

Im letzten Abschnitt haben wir Zeitoperatoren betrachtet, die ausdrucksstärker sind als die bekannten einstelligen Operatoren: Jene lassen sich auf diese nicht zurückführen, umgekehrt aber schon. Die jeweiligen Bedeutungen dieser zweistelligen Tempora konnten wir in Kripke-Strukuren angeben. Dabei fiel Raum für Variationen auf, die wir auch in natürlichsprachlichen Verwendungen von *seit* und *bis* finden.

Nun wollen wir abschließend einige Tempora behandeln, die ebenfalls prominent in der natürlichen Zeitrede vorkommen, z.B. der Ausdruck zur Bezeichnung der Gegenwart, *jetzt*. Dieses und verwandte Tempora fallen dadurch auf, daß sie jenseits der Modellierungsmöglichkeiten von Kripke-Modellen liegen.

Das klingt zunächst merkwürdig. Kann die Gegenwart nicht einfach durch die Abwesenheit eines Tempus ausgedrückt werden? Z.B. so:

$$a \models \text{Es schlägt acht}$$

ist wahr, wenn es *jetzt*, das heißt zum Zeitpunkt a, acht schlägt.

So weit, so gut – oder auch nicht. Was garantiert, daß *jetzt*$=a$ ist? Den Sinne dieser Frage kann man sich auf folgende Weise deutlich machen. Was jetzt wahr ist, daß wird für den jetzigen Zeitpunkt auch in Zukunft wahr bleiben:

> Wenn es *jetzt* acht schlägt, dann wird es in aller Zukunft wahr sein, daß es *jetzt* acht schlägt, und umgekehrt.

Um die formale Strukur eines solchen Satzes wiederzugeben, brauchen wir einen Operator für die Gegenwart, den wir in andere Tempora einbetten können: z.B. "in Zukunft wird es wahr sein (**G**), daß es *jetzt* acht schlägt". Wir notieren diesen *Jetzt*-Operator mit @, und schreiben den Satz so auf: **G**@(Es schlägt acht).

Nun ist es eine Binsenweisheit, daß die Gegenwart durch die Abwesenheit eines Tempus ausgedrückt wird: Wenn es acht schlägt, dann schlägt es *jetzt* acht, und umgekehrt. In einer Formel drücken wir das allgemein so aus:

(J) $$@A \leftrightarrow A.$$

Da (J) nicht zufällig wahr ist, sondern in der Bedeutung des Wörtchens *jetzt* begründet ist ("zur Logik von *jetzt* gehört"), sollten wir die Wahrheitsbedingung für Sätze der Form $@A$ so wählen, daß (J) *gültig* ist. Das ist die *erste Bedingung*. In Modellen auf Kripke-Rahmen können wir diese Bedingung nur so erfüllen:

(A) $$a \models @A \text{ gdw } a \models A.$$

Die *zweite Bedingung* ist, daß A und $@A$ im Skopus eines Tempus, z.B. **P**, nicht einfach austauschbar sind. Wenn es jetzt, zum Zeitpunkt 8, acht ist, dann ist der Satz *Es war einmal sieben* wahr. Aber daraus folgt sicher nicht, daß um 8 der Satz *Es war einmal, daß es jetzt sieben ist* wahr ist – denn jetzt, um 8, ist es acht und nicht sieben. Etwas anders kann man sich das auch klar machen, indem man den – zugegeben, nicht besonders natürlichen aber verständlichen – Satz

> Es war einmal, daß es sieben schlug und jetzt acht schlägt

betrachtet. Diesen Satz, von dem wir einmal annehmen, daß er wahr sei, wollen wir so wiedergeben:

(1) $$\mathsf{P}(7 \wedge @8)$$

Wenn nun 8 und @8 austauschbar sind, dann sagt (1) das Gleiche wie

$$\text{(2)} \qquad \mathsf{P}(7 \wedge 8),$$

d.h.: Es war einmal, daß es zugleich 7 und 8 schlug. Da $7 \neq 8$, kann es einen solchen Zeitpunkt unmöglich gegeben haben. Allgemein gesprochen, sollte also

$$(\dagger) \qquad \mathsf{P}B[A] \leftrightarrow \mathsf{P}B[@A/A]$$

nicht gültig sein.

Es ist unmittelbar klar, daß die beiden genannten Bedingungen, (A) und die Verneinung von (†), nicht zusammenpassen. Hier ist das Argument, welches die Unausweichlichkeit von (†) in Kripke-Modellen, welche (A) erfüllen, zeigt:

$$\begin{aligned} a \models \mathsf{P}@A \text{ gdw } & \exists x : x \prec a \text{ und } x \models @A \quad (\text{nach } (\mathsf{P})) \\ \text{gdw } & \exists x : x \prec a \text{ und } x \models A \quad (\text{nach } (A)) \\ \text{gdw } & a \models \mathsf{P}A \quad (\text{nach } (\mathsf{P})) \end{aligned}$$

Kripke-Modelle sind also grundsätzlich nicht in der Lage, die Bedeutung des *Jetzt*-Operators richtig wiederzugeben. Woran liegt das?

Der *Jetzt*-Operator hat den Zweck, dasjenige was folgt im Hinblick auf den Zeitpunkt der Äußerung zu beurteilen. Wenn ich jetzt, am Punkt a, äußere

Es wird einmal wahr sein, daß es jetzt acht schlägt,

dann frage ich nach einem künftigen Zeitpunkt b zu dem das Urteil richtig sein wird, daß es zum Zeitpunkt der Äußerung, a, acht geschlagen hat. Wir brauchen also *zwei* Referenzpunkte (zwei "Dimensionen") für die Bewertung von *Jetzt*-Aussagen: Den *Zeitpunkt der Äußerung* und den *Zeitpunkt der Beurteilung* (der in der Vergangenheit oder Zukunft liegen kann).

Betrachten wir noch einmal $a \models \mathsf{P}@A$. Der Zeitpunkt der Äußerung, a, legt den Jetzt-Zeitpunkt fest, d.h. $jetzt = a$. Wenn wir die Bedingung für P anwenden, dann erhalten wir

$$\exists x : x \prec a \text{ und } x \models @A.$$

Zwar kommt in dieser Bedingung der Zeitpunkt a der Äußerung noch vor, aber es ist nun der Zeitpunkt x, an dem wir die Wahrheit von $@A$ beurteilen sollen. Eigentlich sollten wir jetzt von x an den Äußerungszeitpunkt a

zurückverwiesen werden, um die Wahrheit von A zu beurteilen. Aber Bedingung (A) fordert fälschlich, daß dafür die Wahrheit von A am Punkt x und nicht *jetzt*, d.h. am Punkt a beurteilt werden soll!

Das Problem lösen wir, indem wir grundsätzlich immer beides, Beurteilungs- und Äußerungspunkt im Auge behalten. Wir schreiben $ab \models A$ und meinen: A ist wahr am Zeitpunkt a so, wie zum Zeitpunkt b geäußert. Wenn sich in einer Formel nichts auf den Äußerungszeitpunkt bezieht, d.h. nur Zeitoperatoren wie **H** und **G**, oder **seit** uns **bis** darin vorkommen, dann brauchen wir die Information aus der zweiten Koordinaten nicht. Aber für die richtige Interpretation des *Jetzt*-Operators ist diese wesentlich. Denn *jetzt* ist ein indexikalisches Tempus, das auf den Zeitpunkt der Äußerung zeigt. Gleichgültig wie tief @ in einen Satz eingebettet sein mag, wir müssen den Zeitpunkt auf den die Äußerung des Satzes zeigt, jederzeit abrufen können. Genau das leistet diese Wahrheitsbedingung:

(@) $$ab \models @A \text{ gdw } bb \models A$$

Die Wahrheitsbedingung ist wie folgt zu lesen:

LR: Wir sind am Punkt a (linke Koordinate) und fragen nach der Wahrheit von $@A$ am Punkt b geäußert (rechte Koordinate). Dazu begeben wir uns nach b (jetzt linke Koordinate) und fragen dort nach der Wahrheit von A.

RL: Der Satz A ist wahr am Punkt b (linke Koordinate) so wie am Punkt b geäußert (rechte Koordinate). Dann ist $@A$, so wie an b geäußert (rechte Koordinate) an einem beliebigen Punkt a (linke Koordinate) wahr.

Wenn wir Sprachen mit dem Jetzt-Operator @ interpretieren wollen, dann müssen wir also zu *zweidimensionalen Interpretationen* übergehen, d.h. wir müssen nun Formeln an Paaren aus $T \times T$ bewerten. Die Interpretationen der nicht-indexikalischen Operatoren bleiben im Kern unverändert. Der Äußerungspunkt bleibt in der Hinterhand nur für den Fall, daß ein *jetzt* davon Gebrauch machen muß. Also:

(P) $$ab \models P \text{ gdw } I(P,a) = 1$$

($\neg$) $$ab \models \neg A \text{ gdw } ab \not\models A$$

($\wedge$) $$ab \models A \wedge B \text{ gdw } ab \models A \;\&\; ab \models B$$

(**H**) $$ab \models \mathbf{H}A \text{ gdw } \forall x : x \prec a \;\Rightarrow\; xb \models A$$

(**G**) $$ab \models \mathbf{G}A \text{ gdw } \forall x : a \prec x \;\Rightarrow\; xb \models A$$

Und die abgeleiteten Bedingungen für P und F:

$$(\mathsf{P}) \qquad ab \models \mathsf{P}A \text{ gdw } \exists x : x \prec a \ \&\ xb \models A$$

$$(\mathsf{F}) \qquad ab \models \mathsf{F}A \text{ gdw } \exists x : a \prec x \ \&\ xb \models A$$

Im Gegensatz zu $\mathsf{P}A$ ergibt sich nun für $\mathsf{P}@A$ diese Wahrheitsbedingung:

$$\begin{aligned} ab \models \mathsf{P}@A \text{ gdw } & \exists x : x \prec a \text{ und } xb \models @A && \text{nach } (\mathsf{P}) \\ \text{gdw } & \exists x : x \prec a \text{ und } bb \models A && \text{nach } (@) \end{aligned}$$

Die letzte Zeile besagt einfach nur, daß A zum Zeitpunkt b wahr sein muß, damit $\mathsf{P}@A$, zum selben Zeitpunkt geäußert, am Punkt a als wahr beurteilt werden kann – plus die Bedingung, daß a überhaupt eine Vergangenheit hat.

Das zuvor problematische Beispiel $\mathsf{P}(7 \wedge @8)$ wird nun so ausbuchstabiert:

$$\begin{aligned} ab \models \mathsf{P}(7 \wedge @8) \text{ gdw } & \exists x : x \prec a \text{ und } xb \models 7 \wedge @8 && \text{nach } (\mathsf{P}) \\ \text{gdw } & \exists x : x \prec a \text{ und } xb \models 7 \ \&\ xb \models @8 && \text{nach } (\wedge) \\ \text{gdw } & \exists x : x \prec a \text{ und } xb \models 7 \ \&\ bb \models 8 && \text{nach } (@) \end{aligned}$$

Die letzte Zeile gibt die Wahrheitsbedingung an, die wir erwarten: Zum Zeitpunkt b geäußert, ist der Satz wahr zum Zeitpunt a, falls es in der Vergangenheit von a einen Punkt gab, zu dem sowohl 7 wahr als auch die Aussage, daß 8 zum Zeitpunkt b wahr ist, wahr war. Etwas vereinfacht sei $a = b$ der jetzige Zeitpunkt. Dann ist der jetzt geäußerte Satz jetzt wahr, wenn 7 einmal wahr war, und es zu diesem vergangen Zeitpunkt auch wahr war, daß jetzt 8 wahr ist.

Und schließlich die Frage nach der ersten Bedingung, (J): Ist $@A \leftrightarrow A$ eine *gültige* Formel? Gültigkeit ist *Wahrheit in allen Modellen* (siehe Def. 1) und eine Formel ist *wahr* in einem Modell, wenn sie an *allen Punkten* des Modells wahr ist. Unsere Interpretationspunkte sind jetzt aber die *Kreuzungspunkte* in der Matrix $T^2 = T \times T$:

$$\begin{array}{cccccc} & a & b & c & \cdots \\ a & aa & ab & ac & \\ b & ba & bb & bc & \\ c & ca & cb & cc & \\ \vdots & & & & \ddots \end{array}$$

Übertragen wir die Definition von Wahrheit in einem Modell auf unseren Fall, dann müssen wir Gültigkeit beurteilen, indem wir das gesamte Feld der Matrix betrachten, also im Sinne von Wahrheit an allen Paaren in allen Modellen:

- A ist *logisch gültig* gdw für alle Modelle $(T, \prec, I)$ und $ab \in T^2 : \ ab \models A$.

Unter dieser Definition ist (J) nicht logisch gültig. Denn sei P ein Atom mit unterschiedlichen Wahrheitswerten an den Punkten a und b. Dann wird (J) am Paar ab falsch:

$$\begin{aligned} ab \models @P \leftrightarrow P \text{ gdw } & ab \models @P \iff ab \models P && \text{nach } (\leftrightarrow) \\ \text{gdw } & bb \models P \iff ab \models P && \text{nach } (@) \\ \text{gdw } & I(P, b) = I(P, a) && \text{nach } (P) \end{aligned}$$

Dennoch können wir in gewisser Weise die erste Bedingung bedienen, daß (J) sich eines besonderen Status erfreue. Dieser Status besteht ja darin, daß wir folgendes *a priori* wissen:

> Wenn die Äußerung zum Zeitpunkt a *Es schlägt acht* zum Zeitpunkt a wahr ist, dann ist die Äußerung zum Zeitpunkt a *Jetzt schlägt's acht* zum Zeitpunkt a wahr, und umgekehrt.

Dieses Zusammenhang gilt nur, wenn wir ausgezeichnete Paare betrachten, nämlich Paare der Form aa auf der Diagonalen der Matrix, bei denen der Punkt der Beurteilung zugleich der Punkt der Äußerung ist. Das *a priori* Wissen, welches durch die Diagonale repräsentiert wird, fassen wir in eine Definition:

- A ist *indexikalisch gültig* gdw für alle Modelle $(T, \prec, v)$ und $a \in T$: $aa \models A$.

Man sieht sofort, daß (J) indexikalisch gültig ist.

Daß wir hier zwei Gültigkeitsbegriffe (und also zwei Begriffen von Äquivalenz) haben, stellt uns in diesem Fall nicht vor die Qual einer Wahl sondern löst vielmehr unser Eingangsproblem, wie die zwei Bedingungen vereinbart werden können. Denn

– einerseits wollen wir eine Äquivalenz im Sinne des Schemas (J), und
– andererseits wollen wir eine Äquivalenz, die Substituierbarkeit in beliebigen Kontexten garantiert.

Indexikalische Äquivalenz gibt uns das erste, logische Äquivalenz das zweite.[7]

[7] Der Ausgangspunkt der neueren Diskussion über zweidimensionale Semantik ist die These, daß Kamps Begriff der indexikalischen Gültigkeit für die Explikation a priorischer Wahrheit genutzt werden kann; siehe [95].

12. Die offene Zukunft

In Michel Houellebecqs Roman *Unterwerfung* wird ein erstaunliches politisches Szenario durchgespielt: eine Regierungsbildung in Frankreich unter Führung einer islamistischen Partei. Darauf mag ein Leser A so reagieren:

> A: Das wird nicht passieren.

Leser B möchte widersprechen. Er sagt aber nicht "Das wird passieren" – so weit möchte B sich nicht hervorwagen –, sondern

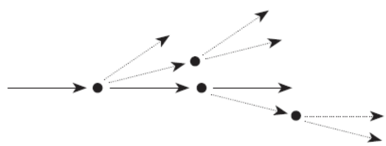
> B: Das kann sehr wohl passieren.

Darauf wieder A:

> A: Ja, aber so wie die Dinge sich entwickeln werden, wird es nicht passieren.

Der Austausch klingt einerseits nach einer Kontroverse über zukünftige Ereignisse; andererseits scheinen die beiden aber auch aneinander vorbei zu reden. A bestreitet nicht die Möglichkeit, auf die B hinweist; A bestreitet, daß diese Möglichkeit wirklich wird.

Beide gehen von derselben Vorstellung aus. Im Gegensatz zur abgeschlossenen Vergangenheit ist die Zukunft noch offen. Die Geschichte kann auf verschiedene Weise weitergehen. Wenn ich mich frage, was die Zukunft bringen wird, dann ist die Frage in gewisser Weise unbestimmt: Welche Zukunft? Die Geschichte ist wie ein Pfad, der an jedem Punkt mehrere mögliche Fortsetzungen offenläßt. Indem wir uns auf dem Pfad vorwärtsbewegen, lassen wir vergangene Möglichkeiten unwiederbringlich hinter uns. Wir können uns das als einen Baum vorstellen:



Ein *Zeitbaum* ist eine strikte Quasiordnung $(T, \prec)$, in der die Vergangenheit eines jeden Punktes eine Kette bildet; d.h. beliebige zwei Punkte haben gemeinsame Vorgänger und $\prec$ ordnet die Vergangenheit jedes Punktes linear:

$$
\begin{gathered}
a \not\prec a, \\
a \prec b \text{ und } b \prec c \;\Rightarrow\; a \prec c; \\
\forall ab \exists c : c \prec a \;\&\; c \prec b, \\
a \prec c \text{ und } b \prec c \;\Rightarrow\; a \prec b \text{ oder } a = b \text{ oder } b \prec a.
\end{gathered}
$$

Ein vollständiger Pfad durch einen Zeitbaum repräsentiert eine mögliche *Geschichte.* Eine Geschichte h ist also eine inklusionsmaximale Kette in T, d.h. wenn $a \in h$ und $a \prec b$ oder $b \prec a$, dann $b \in h$. Die Menge aller Pfade, die durch einen gegeben Punkt a gehen ist die *offene Geschichte* $H(a)$ am Punkt a, also die Vergangenheit von a mit all ihren Verlängerungsmöglichkeiten in die Zukunft: $H(a) = \{h : a \in h\}$.

Es stellt sich nun die Frage nach einer formalen Sprache, welche in der Lage ist, auf alternative Geschichten Bezug zu nehmen. Da die Vergangenheit jedes Punktes eindeutig bestimmt ist, dürfen wir erwarten, daß die Vergangenheitsoperatoren wie zuvor behandelt und Änderungen nur die Zukunftsoperatoren betreffen werden. In den bahnbrechenden Arbeiten von Prior, insbesondere in [232], wurden zwei Sprachen untersucht, die, so Prior, auf Ockham bzw. auf Peirce zurückgeführt werden können. [8]

Peirce-Sprachen. In Zeitbäumen ist der F-Operator offenbar sehr schwach. Eine Formel $\mathsf{F}A$ drückt hier lediglich aus, daß A möglicherweise einmal wahr sein wird. Oft wollen wir aber etwas stärkeres behaupten, nämlich, daß A der Fall sein wird, gleichgültig wie sich die Dinge entwickeln werden. In Peirce-Sprachen dient dazu der Operator $[\mathsf{F}]$ mit der Wahrheitsbedingung

$$[\mathsf{F}] \qquad a \models [\mathsf{F}]A \text{ gdw } \forall h' \in H(a)\ \exists x \in h' \colon a \prec x \text{ und } x \models A.$$

Ein ebenso modalisierter G-Operator würde unter dieser Bedingung stehen:

$$[\mathsf{G}] \qquad a \models [\mathsf{G}]A \text{ gdw } \forall h' \in H(a)\ \forall x \in h' \colon a \prec x \;\Rightarrow\; x \models A.$$

Aber hier ist die Quantifikation über Geschichten ersichtlich redundant, denn nach der Bedingung ist $[\mathsf{G}]A$ am Punkt a genau dann wahr, wenn A an allen Nachfolgern von a wahr ist. Mit anderen Worten, $[\mathsf{G}]$ ist nichts anderes als der schon bekannte Operator G. Nun gilt aber nicht, daß $\neg[\mathsf{F}]\neg A \to \mathsf{G}A$. Daher läßt sich der G-Operator nicht durch Dualisierung aus $[\mathsf{F}]$ definieren; er muß vielmehr als weiterer primitiver Operator in die Sprache eingeführt werden.

Ockham-Sprachen. In Peirce-Sprachen ist die Modalisierung nicht nur in den starken Zukunftsoperator $[\mathsf{F}]$, sondern auch in den schwachen Operator F ($= \neg\mathsf{G}\neg$) eingebaut. Denn F ist nichts anderes als $\langle\mathsf{F}\rangle$ und stellt eine Existenzquantifikation über Geschichten dar. Die Quantifikation über Geschichten ist hier zwar – ähnlich wie im Fall $[\mathsf{G}]$ – redundant, aber sie

[8] Für neuere Untersuchungen siehe insbes. die Darstellung in [117], sowie die Arbeiten von Zanardo.

macht deutlich, daß F auf eine ausgezeichnete, aktuale Geschichte keine Rücksicht nimmt. Zwischen dem so sehr schwach bestimmten F und dem starken $[\mathsf{F}]$ liegt aber offenbar ein Zukunftsoperator, von dem wir immer dann Gebrauch machen, wenn wir etwas über die Zukunft sagen wollen, so "wie sich die Dinge entwickeln werden" (vgl. den Dialog oben). Der einfache F-Operator könnte das leisten, vorausgesetzt, wir fixieren an einem Punkte a eine Geschichte und quantifizieren nur über die Nachfolger von a in *dieser* Geschichte. Für historische Notwendigkeit und Möglichkeit an einem Punkt wollen wir dagegen über die Geschichten quantifizieren, die durch diesen Punkt gehen. Dazu benötigen wir offenbar zwei Parameter, Zeitpunkte und Geschichten, auf die wir bei jeder Beurteilung einer Aussage separat Bezug nehmen können. Mit anderen Worten: Wir brauchen eine Doppelindizierung, wie wir sie schon in der Semantik indexikalischer Ausdrücke kennengelernt haben.

Ockham-Sprachen erweitern eine zeitlogische Sprache mit P und F um einen neuen Operator $\Box$. Formeln beurteilen wir nun an Paaren, bestehend aus einem Zeitpunkt und einer Geschichte, die durch diesen Punkt geht. Im einzelnen sehen die Wahrheitsbedingungen für eine solche Sprache wie folgt aus.

(P) $\qquad ah \models P \text{ gdw } I(P,a) = 1$

$(\neg)$ $\qquad ah \models \neg A \text{ gdw } ah \not\models A$

$(\wedge)$ $\qquad ah \models A \wedge B \text{ gdw } ah \models A \text{ und } ah \models B$

(P) $\qquad ah \models \mathsf{P}A \text{ gdw } \exists x \in h : x \prec a \text{ und } xh \models A$

(F) $\qquad ah \models \mathsf{F}A \text{ gdw } \exists x \in h : a \prec x \text{ und } xh \models A$

$(\Box)$ $\qquad ah \models \Box A \text{ gdw } \forall h' \in H(a) : ah' \models A$

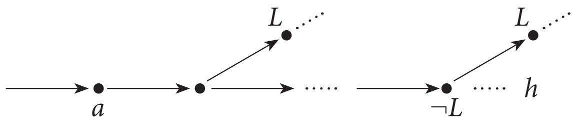
Auf der rechten Seite von (P) ist die Bezugnahme auf h natürlich redundant. In der Bedingung (F) sehen wir dagegen, wie die Geschichtskoordinate h den Quantifikationsbereich für die Zeitkoordinate einschränkt. In der Bedingung für $\Box$ kommt nur die Geschichtskoordinate zum Zuge. Da der Peirce-Operator $[\mathsf{F}]$ sich als $\Box\mathsf{F}$ definieren läßt, umgekehrt aber $\Box$ in Peirce-Sprachen nicht ausgedrückt werden kann, so sind Ockham-Sprachen ausdrucksstärker als Peirce-Sprachen.

Die Positionen von A und B (im Dialog oben) am Paar (a, h) können wir nun mit $\neg\mathsf{F}P$ bzw. $\Diamond\mathsf{F}P$ wiedergeben: B weist darauf hin, daß Houllebecq eine Geschichte h' in $H(a)$ beschreibt in der es einen P-Punkt geben wird. A meint dazu, daß $h' \neq h$ – was B's Äußerung offen gelassen hat.

Auch kompliziertere Betrachtungen über die Zukunft können auf diese Weise dargestellt werden. Hier ist ein Beispiel:

> Zwar wird das Leben auf der Erde schließlich unwiderbringlich untergehen, aber bevor es soweit ist, könnte man zu jedem Zeitpunkt eingreifen, um ewiges Leben zu sichern.

Wir können den Satz im Sinne dieser Formalisierung verstehen (L = Es gibt Leben): $\mathsf{FG}\neg L \wedge \mathsf{G}\Diamond\mathsf{G}L$. An einem Paar (a, h) ist die Formel genau dann wahr, wenn es einen a-Nachfolger in h gibt, ab dem alle Punkte in h nur noch $\neg L$-Punkte sind, zugleich aber von jedem a-Nachfolger x eine Geschichte $h' \in H(x)$ abzweigt, in der alle Punkte L-Punkte sind. Das Diagramm skizziert Modelle, die der Beschreibung genügen:



Gültig ist eine Formel, wenn sie in beliebigen Modellen an allen Paaren wahr ist. Wie spiegelt sich nun die Asymmetrie von abgeschossener Vergangenheit und offener Zukunft in der Gültigkeit von Formeln wieder? Die "fatalistische" (Prior) Formel $\mathsf{F}A \to \Box\mathsf{F}A$ ist ungültig – die Zukunft ist offen. Ein naheliegender Kandidat, die Abgeschlossenheit der Vergangenheit auszudrücken, ist die Formel $\mathsf{P}A \to \Box\mathsf{P}A$. Sie ist jedoch nicht gültig: Man betrachte die Instanz $\mathsf{PF}A \to \Box\mathsf{PF}A$! Das Antezedens schließt nicht aus, daß durch die Vergangenheit eines Beurteilungspunktes nichtaktuale Geschichten gehen, in denen A niemals der Fall sein wird. Allenfalls können wir (gültig) sagen, daß $\mathsf{PF}A \to \Box\mathsf{P}\Diamond\mathsf{F}A$, denn das Antezedens fordert ja die Wahrheit von A in der aktualen Geschichte.[9]

Aus (P) und ($\Box$) folgt unmittelbar, daß für beliebige Atome P, $P \to \Box P$ gültig ist (die Umkehrung gilt für alle Formeln). In gewisser Hinsicht ist das unproblematisch. Denn eine wahre Aussage deren Wahrheit nicht von künftigen Ereignissen abhängt, ist wahr, gleichgültig wie sich die Dinge entwickeln. Aber der Umstand, daß in P kein Zukunftsoperator vorkommt, muß nicht ausschließen, daß P auf andere Weise wesentlich auf die Zukunft Bezug nimmt. Wenn wir diese Möglichkeit einräumen möchten, sollten wir den Beginn der induktiven Definition von $\models$ so modifizieren, daß auch Atome ihre Wahrheitswerte abhängig von einer Geschichte erhalten:

$$(P) \qquad ah \models P \text{ gdw } I(P, ah) = 1.$$

Wir beantworten damit zugleich die Frage, woher der F-Operator den Parameter h nimmt, der seinen Quantifikationsbereich bestimmt.

[9] Das Problem rührt offenbar daher, daß wir in $\mathsf{P}A \to \Box\mathsf{P}A$ für A Formeln einsetzen können, die sich auf die offene Zukunft beziehen. Vgl. dazu die Diskussion in Prior [232, Kap. VII]

III.

MODALLOGIK

1. Kisten und Diamanten

Modallogik wird, in einer ersten Annäherung, meist als die *Logik von Notwendigkeit und Möglichkeit* bezeichnet. Das ist eine sehr enge, mehrdeutige und historisch nicht ganz zutreffende Bestimmung.

Die Bestimmung ist zu *eng*, weil sie z.B. die im vorigen Kapitel behandelte Zeitlogik ausschließt. Wir werden aber bald sehen, daß es wenig Abstraktion bedarf, die Zeitlogik als eine Modallogik unter einer bestimmten Interpretation aufzufassen. Dabei kommt die "rein logische" Theorie ganz gut ohne diese Interpretation aus. Gleiches gilt für deontische, epistemische, dynamische und viele andere Logiken.

Die Bestimmung ist zugleich *mehrdeutig*, denn um was für einen Begriff von Notwendigkeit soll es sich handeln? Die Aussage

$$2+2=4$$

ist eine *arithmetische* Notwendigkeit. Ist die Aussage auch eine logische Notwendigkeit? Frege behauptet dies – aber diese Behauptung markiert eine umstrittene Position (Logizismus) in der Philosophie der Mathematik. Die Aussage

Junggesellen sind unverheiratet

ist wahr aufgrund der Bedeutungen der darin vorkommenden Begriffe; sie ist – wie man sagt – analytisch wahr und als solche eine *analytische* Notwendigkeit. Die Aussage folgt nicht, wie etwa $A \to A$, aus den Gesetzen der Logik und ist in diesem Sinne keine *logische* Notwendigkeit.

Die logische Notwendigkeit von $A \to A$ können wir auf zweierlei Weise einsehen: indem wir die Formel in einem axiomatischen System beweisen oder indem wir zeigen, daß sie in bestimmten Strukturen gültig ist. Logische Notwendigkeit präsentiert sich also in zweierlei Gestalt, als *Beweisbarkeit* und als *Gültigkeit*.

Die Aussage

> Im Vakuum beträgt die Lichtgeschwindigkeit etwa 300 Tausend km in der Sekunde

ist *physikalisch* notwendig. Licht *könnte* sich jedoch auch mit einer anderen Geschwindigkeit bewegen. Das "könnte" deutet hier eine Möglichkeit an, die jedenfalls keine physikalische ist. Dagegen ist

> Wasser ist H_2O

in einem stärkeren Sinne notwendig: Falls die Flüssigkeit im Glas Wasser ist, dann handelt es sich notwendigerweise um H_2O; denn Wasser ist nun einmal identisch mit H_2O und nichts kann etwas anderes sein als das, was es ist – das ist eine *metaphysische* Notwendigkeit.

Vielleicht ist es eine *historische* Notwendigkeit, daß es 1990 zur Wiedervereinigung Deutschlands gekommen ist. So wie die Geschichte bis zu einem bestimmten Zeitpunkt verlaufen ist, war es ab diesem Zeitpunkt vielleicht keine offene Frage mehr, ob Deutschland wieder ein Staat werden würde oder nicht. Im Kapitel II über Zeitlogik haben wir gesehen, daß die Antwort auf die Frage nach der historischen Notwendigkeit eines Ereignisses vom Zeitpunkt der Fragestellung abhängt. Ein nicht-trivialer Determinismus müßte diesen Zeitpunkt weit vor die Abstimmungen über den Einigungsvertrag am 20. September 1990 legen.

Schließlich ist die Bestimmung der Modallogik als Logik der Notwendigkeit auch *historisch* nicht ganz richtig, jedenfalls nicht für die neuere Geschichte der Modallogik. Diese begann mit C.I. Lewis' Versuch einen im Vergleich zur materialen Implikation strikteren Implikationsbegriff zu entwickeln (seit 1912 in einer Reihe von Artikeln; siehe [179, 178] und die kurze Geschichte der Modallogik in [19]). Die strikte Implikation sollte die Vorstellung wiedergeben, daß zwischen Antezedens und Konsequens einer logischen Implikation ein echtes Deduktionsverhältnis besteht. Da ein solches Verhältnis nicht zwischen beliebigen wahren Aussagen bestehen kann, ist die materiale Implikation ungeeignet, logische Implikationen wiederzugeben. (Denn wenn A und B wahr sind, dann impliziert A material B und umgekehrt.) Lewis schlug deshalb Systeme strikter Implikation vor und entwickelte diese axiomatisch (siehe z.B. [177]).

Lewis zufolge sollte eine strikte Implikation derart sein, daß die Wahrheit des Antezedens unverträglich mit der Falschheit des Konsequens ist. Das legt nahe, die strikte Implikation so zu definieren:

$$A \twoheadrightarrow B := \neg\Diamond(A \wedge \neg B).$$

Aber für Lewis stand nicht Möglichkeit oder Notwendigkeit im Zentrum der Theorie sondern Implikation; also war $\twoheadrightarrow$ der grundlegende Operator. Erst später führt er die Verträglichkeit (Konsistenz) zweier Aussagen ein:

$$A \circ B := \neg(A \twoheadrightarrow \neg B).$$

Und erst darauf definiert Lewis Möglichkeit so:

$$\Diamond A := A \circ A \; [= \neg(A \twoheadrightarrow \neg A)].$$

Die Bezeichnungen **S1**, **S2**, **S3**, usw. für Systeme von Modallogiken stammen noch von Lewis. **S1** war so etwas wie ein Basissystem und **S2** war das System, welches er 1912 favorisierte. Die ersten Lewis-Systeme fallen übrigens, wie wir sehen werden, aus dem Rahmen unserer Untersuchung – im wörtlichen Sinne, wie sich zeigen wird, denn es handelt sich dabei nicht um *normale* Modallogiken.

Nach dem bisher Gesagten wird es besser sein, zwischen Modallogik in einem *engeren* und einem *weiteren* Sinne zu unterscheiden. Im engeren Sinne ist das eine Familie logischer Theorien verschiedener Arten von Möglichkeit, Notwendigkeit und davon ableitbarer Begriffe, wie z.B. strikter Implikation. Wenn man diese Theorien mit modernen Mitteln in Angriff nimmt, werden sie schnell unter eine wesentlich allgemeinere Theorie subsumiert. Das ist die Theorie der *Kripke-Rahmen* und einiger ihrer Verallgemeinerungen (ternäre Rahmen, Multirahmen, Nachbarschaftsrahmen, etc.).[1]

In Kripke-Rahmen – oder besser gesagt, in Modellen auf Kripke-Rahmen – gehen wir von der Wahrheit oder Falschheit einer Aussage in verschiedenen Szenarien (Situationen, möglichen Welten) aus. In diesem Kapitel werden wir meist die neutralere Rede von Wahrheitswerten an *Indizes* oder *Punkten* vorziehen. Sodann können wir ganze Bereiche von Punkten betrachten und fragen, für welche Aussagen es Punkte im Bereich gibt, an denen sie wahr sind, oder welche Aussagen an allen Punkten des Bereichs wahr sind. Von einer Aussage, die an mindestens einem Punkt im Bereich wahr ist,

[1] In diesem Kapitel schließen wir uns der in der Modallogik üblichen Rede von "Rahmen" an. Diese sind nichts anderes als das, was wir im Kapitel II über Zeitlogik Strukturen genannt haben.

können wir sagen, daß sie in diesem Bereich *möglich* sei; und eine Aussage, die in dem Bereich nie falsch ist, können wir in einem offensichtlichen Sinne *notwendig* nennen. Mit anderen Worten, die zu erklärende Rede von der Möglichkeit oder Notwendigkeit einer Aussage wird zurückgeführt auf eine einfache Quantifikation über die Punkte eines Bereichs – kurz: modale Operatoren sind Quantoren. Je nachdem, wie wir die Punkte und relevanten Bereiche glossieren, drängen sich die einschlägigen Interpretationen der Operatoren auf: als Zeitmodi, im Sinne verschiedener Arten von Notwendigkeit, als doxastische oder epistemische Einstellungen, oder unter deontischen Lesarten.

In diesem Kapitel gehen wir von der klassischen Aussagenlogik **KL** aus. Unter einer *Modallogik* wollen wir jede Erweiterung von **KL** in einer Sprache mit einem einstelligen Operator $\Box$ ("Kiste") verstehen, der in Kripke-Modellen als Allquantor über Punkte interpretiert wird.[2] Modallogiken sollen wie **KL** unter Modus Ponens und gleichförmiger Einsetzung abgeschlossen sein. Diese Bestimmung ist weit davon entfernt als Definition zu taugen, aber sie lenkt die Gedanken umstandslos in die richtige Richtung. Ein zu $\Box$ dualer Operator $\Diamond$ ("Diamant") ist meist nützlich und so definiert: $\Diamond = \neg\Box\neg$. In Kripke-Modellen wird $\Diamond$ als Existenzquantor über Punkte interpretiert und repräsentiert so Mögichkeit (im weitesten Sinne). In allen bekannten Modallogiken sind $\Box$ und $\Diamond$ wechselseitig definierbar.

Es kann keinen vernünftigen Zweifel daran geben, daß Kripke-Rahmen eine Theorie abgeben, in der sich wichtige Fragen über Modallogik stellen und beantworten lassen. Zu diesen Fragen gehört vornehmlich die nach den *logischen* Eigenschaften verschiedener Begriffe von Notwendigkeit und Möglichkeit. Aber taugen Kripke-Rahmen auch als Gerüst für eine philosophische Theorie z.B. metaphysischer Notwendigkeit? Das ist eine durchaus kontroverse Frage, deren Beantwortung im wesentlichen davon abhängt, welche Erklärungen wir von einer solchen Theorie erwarten und welchen ontologischen Status die theoretischen Termini der Theorie – Punkt ("mögliche Welt") und zweistellige Relation ("Erreichbarkeit") – plausiblerweise haben können. Die Sache verhält sich nicht anders als mit der Mengenlehre oder der Arithmetik. Auch diese sind Theorien, die über Objekte quantifizieren – Mengen und Zahlen – deren ontologischer Status aus philosophischer Perspektive unklar sein mag. Die Arithmetik ersetzt keine philosophische Auseinandersetzung mit dem Zahlbegriff (wie etwa in Freges [79]). Aber sie gibt eine Charakterisierung des Gegenstands der philosophischen Untersuchung vor. Ebensowenig ist die Beschreibung der Kripke-Rahmen *per*

[2] Die Ausgangsbasis **KL** ist nicht zwingend: Auch andere Logiken, wie intuitionistische oder relevante Logiken lassen sich modallogisch erweitern.

se eine philosophische Theorie metaphysischer Notwendigkeit (wie etwa in Lewis' [181]). Aber Kripke-Rahmen bestimmen den Gegenstand einer solchen Theorie; vgl. dazu auch die folgende Anmerkung. Insofern Notwendigkeit einer der zentralen Begriffe der Philosophie ist, ist die Theorie der Kripke-Rahmen für den Philosophen (und nicht nur für diesen), was Mengenlehre und Arithmetik für den Mathematiker (und nicht nur für diesen) sind.

Anmerkung zur Geschichte der modalen Semantik. Die Allgegenwart der Begriffe 'Möglichkeit' und 'Notwendigkeit' in der philosophischen Theoriebildung seit der Antike steht in einem merkwürdigen Gegensatz zu der Unbekümmertheit, mit der diese Begriffe bis ins späte 19. Jahrhundert verwendet wurden. Meist wurde unter 'Notwendigeit' kaum mehr als die Abwesenheit von Zufall oder ein gewisser gedanklicher Zwang verstanden.

Auch in den Arbeiten von C.I. Lewis zur Modallogik findet sich kaum ein Versuch einer inhaltlichen Klärung. Die Vielzahl modaler Logiken, die C.I. Lewis vorstellt, machen eine solche Klärung aber umso drängender. Es wird offensichtlich ein Kriterium benötigt, um zwischen diesen Systemen auszuwählen. Das kann aber nichts anderes heißen als: Welches System der Modallogik ist für ein gegebenes Verständnis der Modalitäten das richtige? Wir brauchen also ein Verständnis von Modalitäten, das von axiomatischen Darstellungen zunächst unabhängig ist aber einen vergleichbaren Präzisionsgrad hat.

Gödel (1933) [112] war vermutlich der erste, der dieses Problem in eine Form brachte, die eine klare Beantwortung ermöglichte. Er interpretierte Notwendigkeit als *Beweisbarkeit* in einer bestimmten Art von axiomatischen Systemen; wir werden darauf zurückkommen (p. 187ff.). In den 40er Jahren entwickelte Carnap [45, 46] einen dazu komplementären, auf Modellen der klassischen Aussagenlogik beruhenden Ansatz. Danach wird Notwendigkeit als logische *Gültigkeit*, d.h. Wahrheit in allen Modellen interpretiert. Diesen Ansatz werden wir im nächsten Abschnitt weiterverfolgen, wo er uns in der Klausel ($\Box$) auf p. 84 entgegentritt.

In den frühen 50er Jahren begannen Kanger [159] und Hintikka [137] auf der Grundlage von Carnaps Arbeiten auch andere Deutungen des Notwendigkeitsoperators in den Blick zu nehmen, insbesondere epistemische, doxastische und deontische Interpretationen. Wenn $\Box A$ beispielsweise ausdrücken soll, daß in einer bestimmten Situation a eine Person glaubt, daß A der Fall ist, dann können wir das so verstehen, daß in allen Situation, die er für möglich hält, A wahr ist. Wir quantifizieren hier nicht über alle Möglichkeiten schlechthin, sondern nur über diejenigen Möglichkeiten, welche die Person in a für möglich, d.h. für nicht ausgeschlossen ("erreichbar") hält. Die so postulierte Relation der Erreichbarkeit erweist sich als Mittel, eine große Spannbreite von Modalitäten wiederzugeben. Auch die in den 50er Jahren entstandenen Arbeiten von Prior [231, 232] zur Zeitlogik dürfen in diesem Zusammenhang nicht unerwähnt bleiben. Prior bezog zeitlogische Axiome auf Mengen von Zuständen, die durch eine Relation zeitlicher Erreichbarkeit ("früher als") angeordnet sind. Verschiedene Arten zeitlicher Modalitäten unterschied er so aufgrund von Korrespondenzen zwischen Axiomen einer modalen Sprache einerseits und Eigenschaften der Ordnungsrelation andererseits. Die zeitlogische Interpretation war Gegenstand des vorangegangenen Kapitels.

Vermutlich unabhängig von Kanger, Hintikka und Prior hat Kripke (1959) [164] im Prinzip dieselbe Semantik gefunden. Viele Details dieser beinahe simultanen Entdeckung sind unbekannt. Ein wesentlicher Anstoß für Kripkes Arbeiten könnte ein Repräsentationsresultat in Jónsson und Tarski (1952) [158] gewesen sein. In deren zwei Aufsätzen über Boole'sche Algebren mit Operatoren ist zwar von Modallogik nicht die Rede, aber die Anwendung ist, zumindest im Rückblick, recht offensichtlich. Das Repräsentationsresultat zeigt, wie aus "modalen Algebren" (ein späterer Terminus von E.J. Lemmon [170]) mengentheoretische Modelle abgeleitet werden können, die wir heute als Kripke-Modelle kennen.[3]

Aufbauend auf der Grundidee von Carnap, ist die Modelltheorie der Modallogik, so wie wir sie heute kennen, in einem Zeitraum weniger Jahre entstanden. Die notwendige Verallgemeinerung der Grundidee scheint als stabiler Fluchtpunkt von mehreren Logikern etwa gleichzeitig und unabhängig voneinander gefunden worden zu sein. Was aber drängte dazu, Mitte der 50er Jahre verschiedene Theorieelemente zu einer Semantik der Modallogik zusammenzuführen? Eine Antwort auf diese Frage kann naturgemäß nur spekulativ ausfallen. Einen wesentlichen Anteil dürften Quines Angriffe auf die Möglichkeit einer philosophisch sauberen Theorie der Modalitäten gehabt haben.[4] Die Modelle der Modallogik zeigen, wie sich intensionale Aussagenoperationen auf Quantifikation über geeignete Objekte zurückführen und somit besser verstehen lassen. Quine hat diese Leistung anerkannt, jedoch für letztlich unzureichend gehalten. Denn er fürchtete, daß sich die extensionale Semantik der modalen Aussagenlogik nicht zu einer

---

[3] Ausführlichere Darstellungen zur Geschichte der sogenannten Kripke-Modelle sind in [188] und [31, Kap. 1] enthalten. Eine aus systematischer Perspektive rekonstruierte Geschichte bietet Humberstone in [144, §2.10] an. Zum Verhältnis von mengentheoretischer und algebraischer Semantik für die Modallogik siehe insbesondere Goldblatt [114, Kap. 1], und [31, Kap. 5].

[4] Siehe z.B. "The problem of interpreting modal logic" (1947) [233] und weitere Aufsätze, zum Teil abgedruckt in [235]. In diesem Zusammenhang sei auch auf die aufschlußreiche Korrespondenz zwischen Carnap und Quine verwiesen [238].

Semantik der modalen Prädikatenlogik erweitern läßt, ohne fragwürdige metaphysische Positionen über Gegenstände und ihre Eigenschaften einzunehmen. Da modale Prädikatenlogik nicht Gegenstand dieses Buches ist, dürfen wir es bei dem Hinweis belassen.

2. Die Logik logischer Notwendigkeit

In der nicht-modalen Aussagenlogik kam es uns wesentlich darauf an, den Begriff der *logischen Wahrheit* (Gültigkeit) zu untersuchen. Gültig zu sein, ist eine Eigenschaft, die Sätze haben können. Sie ist grammatisch nicht unähnlich der Eigenschaft (in einem genügend weiten Sinne) negiert zu sein, mit einem anderen Satz in Konjunktion zu treten oder als Antezedens einer Implikation aufzutreten. Es ist daher naheliegend unsere nicht-modale Sprache $\mathcal{L}$ so zu erweitern, daß wir darin ausdrücken können, daß ein Satz die Eigenschaft hat, gültig zu sein. Das ist der erste Schritt, um logische Gültigkeit selbst zum Gegenstand einer logischen Theorie zu machen.

Den zweiten Schritt finden wir in Carnap [46, §39] so beschrieben:

> Wir erkennen eine starke Ähnlichkeit zwischen zwei *explicanda*, nämlich der logischen Notwendigkeit einer Aussage und der logischen Wahrheit eines Satzes. Nun besitzen wir für letzteren Begriff ein exaktes *explicatum* im semantischen Begriff der L-Wahrheit, welche definiert ist auf der Grundlage der Begriffe einer Zustandsbeschreibung und eines Bereichs. Es scheint mir daher der natürlichste Weg zu sein, als *explicatum* für logische Notwendigkeit diejenige Eigenschaft einer Aussage zu nehmen, welche der L-Wahrheit eines Satzes entspricht. [46, p. 174; Übersetzung A.F.]

Carnaps L-Wahrheit ist das, was wir in Rahmen der klassischen Aussagenlogik als logische Wahrheit definiert haben. Wir haben gesagt:

> Ein Satz A ist gd logisch wahr oder gültig (*logisch notwendig*, wie wir fortan sagen werden), wenn A in allen Modellen wahr ist.

Modelle sind im wesentlichen Carnaps "Zustandsbeschreibungen". Nun unterscheiden sich die Modelle $(\{0,1\}, I)$ der klassischen Aussagenlogik einzig dadurch, wie sie die atomaren Formeln bewerten (interpretieren), d.h. die Wahrheitswerte 0 und 1 mittels der Funktion I über die Atome verteilen. Den Rest bestimmen die Wahrheitstafeln für alle Modelle auf die gleiche Weise. Also können wir auch sagen:

> Ein Satz A ist gd logisch notwendig, wenn er unter allen Interpretationen wahr ist.

Diese einfache Idee werden wir gleich systematisch ausbuchstabieren.

Aus einer Bewertung I der Atome und den bekannten Wahrheitstafeln ergibt sich auf natürliche Weise eine *Wahrmacher-* oder *Erfüllungsrelation*, $\models$ zwischen der Bewertung I und beliebigen Formeln (für $I \models A$ lesen wir: "I macht A wahr"):

$$(P) \qquad I \models P \text{ gdw } I(P) = 1$$

$$(\neg) \qquad I \models \neg A \text{ gdw } I \not\models A$$

$$(\wedge) \qquad I \models A \wedge B \text{ gdw } I \models A \;\&\; I \models B$$

Jetzt wollen wir einen weiteren einstelligen Junktor $\Box$ einführen so, daß $\Box A$ bedeuten soll: A ist logisch notwendig. Wann macht eine Bewertung I eine Formel A logisch notwendig? Wir haben es gerade gesagt: wenn A unter *allen* Bewertungen wahr ist:

$$(\Box) \qquad I \models \Box A \text{ gdw für alle } I' : I' \models A$$

In dieser Wahrheitsbedingung quantifizieren wir also über einen Bereich von Interpretationen I, wie sie die klassischen Modelle $(\{0,1\}, I)$ charakterisieren. D.h. für die Interpretation einer Sprache $\mathcal{L}^{\Box}$, in der $\Box$ für logische Notwendigkeit stehen soll, genügt es nicht nur eine Interpretation zu betrachten, sondern wir müssen über *alle* Interpretationen quantifizieren können:

$$(\{0,1\}, I, I', I'', ...).$$

Im Grunde also betrachten wir einen Bereich von Interpretationen und fragen nach der Wahrheit von Formeln relativ zu einer Interpretation. Die übliche Art, das aufzuschreiben, geht so:

· Wir erwähnen fortan nicht mehr die Menge der Wahrheitswerte $\{0,1\}$. Sofern nicht Gegenteiliges vereinbart wird, gehen wir davon aus, daß es nur diese zwei Wahrheitswerte gibt.

· Statt die potentielle Unendlichkeit von Bewertungen mit Strichen und Punkten anzudeuten, indizieren wir sie mit einer Indexmenge W. Oder, was auf das Gleiche hinausläuft: Wir verwenden den Buchstaben I jetzt für eine Funktion, die sich einen Index aus W greift und, so relativiert, Atome bewertet:

$$(W, I) \text{ mit } I : \text{ATM} \times W \longrightarrow \{0,1\}.$$

Als Variablen für Indizes benutzen wir $a, b, c, ...$ Diese Indizes nennen wir fortan *Punkte*, etwas phantasievoller auch *mögliche Welten*.

· In der Angabe der Wahrheitsbedingungen lassen wir jetzt einfach die Punkte (Indizes) für die korrespondierenden Bewertungen stehen.

Danach sehen die Wahrheitsbedingungen in (W, I) so aus:

$$(P) \qquad a \models P \text{ gdw } I(P, a) = 1$$

$$(\neg) \qquad a \models \neg A \text{ gdw } a \not\models A$$

$$(\wedge) \qquad a \models A \wedge B \text{ gdw } a \models A \ \& \ a \models B$$

$$(\Box) \qquad a \models \Box A \text{ gdw } \forall b \in W : b \models A$$

- Ein Satz ist *wahr in einem Modell* (W, I), wenn ihn alle Punkte wahr machen.

In einem bestimmten Modell (W, I) repräsentiert W eine nichtleere Menge möglicher Bewertungen der Atome. Es mag durchaus sein, daß ein Modell nicht alle möglichen Bewertungen umfaßt (siehe aber die folgende Anmerkung!). Vielleicht besteht es aus nur *einem* Punkt a und nimmt so nur eine einzige Bewertung in den Blick. In diesem Fall ist $\Box A$ wahr im Modell, wenn $a \models A$: Dieses Modell ist halt so, daß der Bereich der zu berücksichtigenden Bewertungen aus nur einem Punkt besteht. Damit eine Formel logisch wahr (gültig) ist, müssen wir von solchen Besonderheiten einzelner Modelle absehen. So kommen wir zur folgenden

DEFINITION 1.
1. A ist *wahr* in $\mathcal{M} = (W, I)$ gdw $\forall a \in W : a \models A$. Notation: $\mathcal{M} \models A$.
2. A ist *gültig* (in der Klasse $\mathbb{M}$ von Modellen) gdw $\forall \mathcal{M}$ $(\in \mathbb{M}) : \mathcal{M} \models A$. Notation: $\models A$ bzw. $\mathbb{M} \models A$.

Anmerkung. Wir haben hier Carnaps Ansatz in die Richtung von Kripkes Theorie gebürstet. Aber in Carnaps Theorie ist W nicht eine *beliebige* (nichtleere) Menge von Zustandsbeschreibungen. Vielmehr ist W die Menge *aller* Zustandsbeschreibungen (Interpretationen der Atome in den Wahrheitswerten 0 und 1). Gültigkeit im Sinne Carnaps ist Wahrheit in diesem "totalen" Modell. Nun ist es aber so, daß es im totalen Modell für jedes Atom P eine Zustandsbeschreibung gibt, in der P falsch ist. Aber dann ist $\Box P$ an allen Punkten falsch, d.h. $\neg\Box P$ ist wahr im totalen Modell und also eine gültige Formel, für jedes Atom P. Man mag es als logische Wahrheit akzeptieren, daß Atome nicht logisch wahr sind. Aber $\neg\Box P$ ist kein Theorem von **S5**, Carnaps Zielsystem. Carnap beschreibt seine Lösung in [46]; vgl. auch [191], [281], und [259].

3. Was wir über logische Notwendigkeit wissen

Aufgrund der Definition können wir schon einiges über den Begriff der logischen Notwendigkeit – so, wie er sich in Modellen (W, I) darstellt – herausfinden.

1. An jedem Punkt werden die Operatoren $\neg$ und $\wedge$ klassisch, d.h. nach den Wahrheitstafeln, behandelt. Alle Sätze, die aufgrund wahrheitsfunktionaler Zusammensetzung gültig sind, werden also an jedem Punkt wahr und sind damit gültig. Die Logik logischer Notwendigkeit ist eine *Erweiterung* der klassischen Logik.
2. Aus 1 folgt, daß alle Punkte unter *Modus Ponens* abgeschlossen sind: Wenn sowohl A als auch $A \to B$ an einem Punkt wahr sind, dann ist auch B an diesem Punkt wahr. Es gilt dann ferner, daß wenn A und $A \to B$ an allen Punkten eines beliebigen Modells wahr sind, dann ist auch B an allen Punkten eines solchen beliebig gewählten Modells wahr. D.h. auch die Eigenschaft der Gültigkeit steht unter der Regel MP.
3. Wenn A gültig ist, dann ist A in allen Modellen an jedem Punkt wahr. Wenn aber A an jedem Punkt wahr ist, dann ist auch $\Box A$ an jedem Punkt wahr. Also gilt: Wenn A gültig ist, dann ist auch $\Box A$ gültig. Das ist die *Notwendigkeitsregel*.
4. Wenn $A \to B$ an einem Punkt a notwendig ist, dann ist $A \to B$ an allen Punkten wahr. Wenn nun auch A notwendig, also an allen Punkten wahr ist, dann muß B ebenfalls an allen Punkten wahr, d.h. notwendig sein. Das gilt für beliebige Punkte a in beliebigen Modellen; kurzum, die Formel $\Box(A \to B) \to (\Box A \to \Box B)$ ist gültig.
5. Logische Notwendigkeit an einem Punkt impliziert immer Wahrheit an diesem Punkt. Also ist $\Box A \to A$ eine gültige Formel.
6. Wenn etwas logisch notwendig ist, dann ist diese Tatsache selbst logisch notwendig: $\Box A \to \Box\Box A$. Denn angenommen, A sei an allen Punkten wahr; dann ist es gleichgültig, welchen Punkt wir betrachten: Überall ist $\Box A$ wahr. So dürfen wir mit $(\Box)$ von rechts nach links schließen, daß $\Box\Box A$ am Ausgangspunkt unserer Überlegung wahr ist.
7. Was wahr ist, muß $(\Box)$ möglich $(\Diamond)$ sein. Das drückt diese Formel aus: $A \to \Box\Diamond A$. Angenommen A sei wahr am Punkt a. Dann ist es an jedem $(\Box!)$ Punkt b wahr, daß es einen Punkt gibt $(\Diamond!)$ – nämlich a –, an dem A wahr ist. Die Formel ist also gültig.
8. Was möglich ist, das ist notwendigerweise möglich: $\Diamond A \to \Box\Diamond A$. Wir sind an einem Punkt a und stellen fest, daß es einen Punkt gibt $(\Diamond)$,

an dem A wahr ist. Aber diese Feststellung können wir von jedem beliebigen ($\Box$!) Punkt aus machen, und das ist selbst etwas, was am Punkt a wahr ist. Aus der Wahrheit von $\Diamond A$ an einem Punkt folgt so auch die Wahrheit von $\Box\Diamond A$ am selben Punkt.

4. Das System S5

Was wir soeben über logische Notwendigkeit herausgefunden haben, wollen wir als ein axiomatisches System aufschreiben:

τ	Jede Tautologie
K	$\Box(A \to B) \to (\Box A \to \Box B)$
T	$\Box A \to A$
4	$\Box A \to \Box\Box A$
B	$A \to \Box\Diamond A$
5	$\Diamond A \to \Box\Diamond A$
Regeln	$\dfrac{A \quad A \to B}{B}$ MP $\qquad \dfrac{A}{\Box A}$ RN

Die Begriffe einer *Ableitung* und eines *Theorems* im Hinblick auf **S5** sind wie üblich definiert. Die Bezeichnungen der logischen Prinzipien folgt hier und im Folgenden einer Systematisierung üblicher Namen, die von Chellas eingeführt wurde.

- Die Bezeichnungsregeln nach Chellas [53] sind einfach. Sei X $= A \to B$ ein Formelschema; dann ist X_c die Umkehrung $B \to A$ von X, und X! ist die Äquivalenz $A \leftrightarrow B$. Wenn X ein Schema “über $\Box$” ist, dann ist X$\Diamond$ das duale (siehe unten) Schema “über $\Diamond$”.

Zu beachten ist, daß das klassische *Deduktionstheorem*, d.h.

$$\text{wenn } X, A \vdash B, \text{ dann } X \vdash A \to B,$$

in der Gegenwart der Regel RN nicht mehr gelten kann. Denn anderenfalls könnten wir so schließen:

1	(1)	A	Annahme
1	(2)	$\Box A$	(1) RN
	(3)	$A \to \Box A$	Deduktionstheorem

$A \to \Box A$ ist *per se* schon keine gute Idee. Zusammen mit T hätten wir dann die trivialisierende Äquivalenz $A \leftrightarrow \Box A$. Auch unsere Modelle sprechen gegen das Schema $A \to \Box A$. (Man betrachte die Instanz $P \to \Box P$ in einem Modell mit zwei Punkten a und b so, daß $a \models P$ und $b \models \neg P$.) Es dürfte klar sein, daß das Deduktionstheorem weiterhin für Ableitungen gilt, in denen von der Regel RN kein Gebrauch gemacht wird. Und *falls* die Regel auf eine Annahme A angewandt wird, um daraus $\Box A$ zu erzeugen, dann können wir das Deduktionstheorem offenbar schadlos anwenden, wenn die Annahme A selbst von der Form $\Box A$ ist; denn in diesem Fall generieren wir eine Instanz von 4. Wir dürfen also erwarten, daß in der Modallogik eingeschränkte Varianten des Deduktionstheorems gelten. Wir werden darauf zurückkommen (siehe p. 107).

Offensichtlich ist das Verhalten der Kiste $\Box$ dem der starken Tempus-Operatoren H (immer in der Vergangenheit) und G (immer in der Zukunft) recht ähnlich. Beinahe alle logischen Prinzipien in der Axiomatisierung von **S5** haben wir dort schon angetroffen. Neu sind nur die Axiome T und 5.

· Das Schema T ist unter der jetzigen Interpretation zwingend. Wenn wir die Kiste jedoch im Sinne von H (immer in der Vergangenheit) interpretieren, dann drückt T, also $\mathsf{H}A \to A$, etwas Falsches aus: Alles bleibt, wie es immer war. Ähnliches gilt für die Zukunftsinterpretation G (immer in der Zukunft).

· Auch das Schema 5 hat in der Zeitlogik keinen Platz. Man betrachte die G-Entsprechung

$$\mathsf{F}A \to \mathsf{GF}A.$$

Es besagt soviel wie: Wenn irgendwann einmal A, dann wird es immer der Fall sein, daß irgendwann einmal A. (Wenn ich irgendwann einmal im Lotto gewinnen werde, dann wird es immer so sein, daß ich irgendwann einmal im Lotto gewinne.) Man kann sich schnell davon überzeugen, daß dieses Schema in zeitlogischen Rahmen nicht gültig sein kann. Im hier skizzierten Gegenbeispiel wird P zum Zeitpunkt b wahr, danach nie wieder:

$$\begin{array}{ccccc} a\bullet & \longrightarrow & b\bullet & \longrightarrow & c\bullet\cdots \\ \mathsf{F}P & & P & & \neg P\cdots \\ \neg\mathsf{GF}P & & \neg\mathsf{F}P & & \end{array}$$

Aus den Axiomen und Regeln für **S5** können wir einige weitere wichtige logischen Prinzipien ableiten. So gelten zum Beispiel die folgenden Regeln der

Monotonie $$\frac{A \to B}{\Box A \to \Box B}\ \text{RM} \qquad \frac{A \to B}{\Diamond A \to \Diamond B}\ \text{RM}\Diamond.$$

Diese Regeln lassen sich leicht ableiten:

$$\dfrac{\dfrac{A \to B}{\Box(A \to B)}\ \text{RN} \qquad \text{K}}{\Box A \to \Box B}\ \text{MP} \qquad \dfrac{\dfrac{\dfrac{A \to B}{\neg B \to \neg A}\ \textbf{KL}}{\Box\neg B \to \Box\neg A}\ \text{RM}}{\neg\Box\neg A \to \neg\Box\neg B}\ \textbf{KL}$$

Aus den Monotonieregeln folgen unmittelbar die Regeln der

Kongruenz $$\frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B}\ \text{RE} \qquad \frac{A \leftrightarrow B}{\Diamond A \leftrightarrow \Diamond B}\ \text{RE}\Diamond.$$

Ferner gilt die Regel RK der

Normalität $$\frac{A_1 \wedge \cdots \wedge A_n \to B}{\Box A_1 \wedge \cdots \wedge \Box A_n \to \Box B} \quad n \geq 0.$$

Diese Regel leiten wir durch Induktion über n aus RN und K her.

Die Induktionsbasis $n = 0$ reduziert RK auf RN. Sodann gehen wir von der Induktionshypothese

IA $$\frac{A_1 \wedge \cdots \wedge A_{n-1} \to B}{\Box A_1 \wedge \cdots \wedge \Box A_{n-1} \to \Box B}$$

aus und nehmen ferner an, daß $A_1 \wedge \cdots \wedge A_n \to B$. Daraus folgt durch $\to$-Exportation $A_1 \wedge \cdots \wedge A_{n-1} \to (A_n \to B)$. Darauf wenden wir die IA an und erhalten $(*)$ $\Box A_1 \wedge \cdots \wedge \Box A_{n-1} \to \Box(A_n \to B)$. Nun impliziert nach K das Konsequens von $(*)$ die Formel $\Box A_n \to \Box B$. Transitivität und Importation für $\to$ ergibt dann das Gewünschte: $\Box A_1 \wedge \cdots \wedge \Box A_n \to \Box B$.

Der Fall $n = 0$ von RK ist RN; der Fall $n = 1$ ist RM; im Fall $n = 2$ ist RK die Regel der

Regularität $$\frac{A \wedge B \to C}{\Box A \wedge \Box B \to \Box C}\ \text{RR}.$$

Das offensichtlich strikte Implikationsverhältnis zwischen den besprochenen vier Regeln stellt sich also so dar:

$$\text{RK} \ \Rightarrow\ \text{RR} \ \Rightarrow\ \text{RM} \ \Rightarrow\ \text{RE}.$$

Eine Modallogik ist

- *kongruent*, wenn sie unter RE abgeschlossen ist,
- *monoton*, wenn sie unter RM abgeschlossen ist,
- *regulär*, wenn sie unter RR abgeschlossen ist, und
- *normal*, wenn sie unter RK abgeschlossen ist.

Die in diesem Kapitel zu besprechenden Modallogiken sind alle normal. Um schwächere als normale Logiken zu modellieren, müßten wir andere als die hier betrachteten Kripke-Rahmen betrachten. Sogenannte Nachbarschaftsrahmen, die wir am Schluß des Kapitels kurz betrachten werden, reichen bis zu den bloß kongruenten Modallogiken.

Ein weiteres wichtiges Prinzip ist die Regel der

Dualität $$\frac{A \to B}{B^d \to A^d}\ \text{Dual.}$$

Hier bezeichnet A^d die zu A duale Formel. Sie entsteht aus A, wenn alle Atome durch ihre Negationen ersetzt, sowie $\vee$ gegen $\wedge$ sowie $\Diamond$ gegen $\Box$ vertauscht werden. Für die Junktorenmenge, mit der wir üblicherweise arbeiten, können wir die Dualfunktion so definieren:

$$\begin{aligned}
P^d &= \neg P\\
\bot^d &= \top\\
(\neg A)^d &= \neg A^d\\
(A \wedge B)^d &= A^d \vee B^d\\
(A \vee B)^d &= A^d \wedge A^d\\
(A \to B)^d &= \neg A^d \wedge B^d\\
(\Box A)^d &= \Diamond A^d\\
(\Diamond A)^d &= \Box A^d
\end{aligned}$$

Es ist klar, daß $A^{dd} = A$, so daß auch die Umkehrung der Dualitätsregel gilt. Wir haben die Dualitätsregel im Kapitel II über Zeitlogik bereits bewiesen. Die Durchsicht des Beweises zeigt, daß die Regel für jede Modallogik gilt, in der äquivalente Formeln im Skopus von $\Box$ und $\Diamond$ ausgetauscht werden dürfen, also für jede Modallogik, die unter der Kongruenzregel RE abgeschlossen ist. Wir benutzen die Regel hier als Rechtfertigung, bestimmte Schemata zu identifizieren. D.h.

$$\Box A \to A \text{ und } A \to \Diamond A$$

oder

$$\Diamond A \to \Box\Diamond A \text{ und } \Diamond\Box A \to \Box A$$

können wir als im wesentlichen identische Schemata betrachten – "identisch mod d". Sinngemäß gilt das ebenso für Regelpaare wie RE und RE$\Diamond$.

Bei der Betrachtung des Duals von T,

T$\Diamond$ $$A \to \Diamond A,$$

stellt sich schnell heraus, daß die obige Axiomatisierung von **S5** redundant ist. Denn das Axiom B folgt unmittelbar aus 5 und T$\Diamond$.

Schließlich wollen wir in einem Satz ein Ergebnis festhalten, daß im wesentlichen schon im Abschnitt 3 vorlag.

SATZ 2. (RICHTIGKEIT) *Das System* **S5** *ist richtig für die oben beschriebenen Modelle (W,I): Wenn* $\vdash A$ *in* **S5**, *dann* $\models A$ *in allen Modellen.*

BEWEIS. Induktion über die Länge einer Ableitung in **S5**. Dazu müssen wir zeigen, daß alle Axiome gültig sind (Induktionsbasis) und daß die zwei Regeln MP und RN die Eigenschaft der Gültigkeit von den Prämissen an die Konklusion weitergeben (Induktionsschritt). Etwas informell haben wir das bereits im Abschnitt 3 getan. Wir führen hier die detaillierte Verfizierung für die Axiome 5 und K vor.

Ad 5: An einem beliebigen Punkt a ist zu zeigen (zz), daß $a \models \Diamond A \to \Box\Diamond A$. Wir nehmen also an: $a \models \Diamond A$, d.h. $\exists b : b \models A$. Es bleibt zz, daß $a \models \Box\Diamond A$, d.h. $\forall x : x \models \Diamond A$, d.h. $\forall x \exists y :\models A$ – was durch die Annahme gegeben ist.

Ad K: Zz: $a \models \Box(A \to B) \to (\Box A \to \Box B)$; d.h. wenn $a \models \Box(A \to B)$, dann $a \models \Box A \to \Box B$; d.h. wenn $a \models \Box(A \to B)$ und $a \models \Box A$, dann $a \models \Box B$.

1	(1)	$a \models \Box(A \to B)$	Annahme
2	(2)	$a \models \Box A$	Annahme /zz $a \models \Box B$
1	(3)	$\forall x : x \models A \to B$	1 ($\Box$)
1	(4)	$\forall x :$ wenn $x \models A$, dann $x \models B$	3 ($\to$)
2	(5)	$\forall x : x \models A$	2 ($\Box$)
1,2	(6)	$\forall x : x \models B$	4,5
1,2	(7)	$a \models \Box B$	6 ($\Box$)

∎

Wir wissen nun, daß jede in **S5** ableitbare Formel gültig ist. Wie steht es um die umgekehrte Behauptung, die Vollständigkeit: Ist jede gültige Formel ein Theorem von **S5**? Wenn das so ist, dann axiomatisiert **S5** genau die Logik logischer Notwendigkeit —jedenfalls in dem Sinne, in dem wir diesen Begriff in den Modellen der Art (W, I) eingefangen haben.

Korollar 3. *(Konsistenz)*
*Jede Formelmenge $L \subseteq$ **S5** ist widerspruchsfrei.*

Beweis. Angenommen L enthält Formeln A und $\neg A$ (äquivalent: $L \vdash \bot$). Dann ist A und $\neg A$ in **S5** und also, nach dem gerade bewiesenen Satz, $\models A$ und $\models \neg A$ in allen Modellen. Aber das bedeutet, daß für alle Punkte a eines Modells, $a \models A$ und $a \models \neg A$ – im Widerspruch zur Modellbedingung ($\neg$)! ∎

Satz 4. (Vollständigkeit) *Das System* **S5** *ist vollständig für die oben beschriebenen Modelle* (W, I)*: Wenn* $\models A$ *in allen Modellen, dann* $\vdash A$ *in* **S5**.

Wir werden diesen Satz erst später beweisen. Zuvor wollen wir Modelle betrachten, die in der Lage sind, auch andere Begriffe von Notwendigkeit wiederzugeben. Dabei werden wir auf eine *minimale Modallogik* treffen, die uns als Ausgangspunkt der weiteren Untersuchungen dienen wird. Dies ist die Logik, welche die klassische Logik nur um die Regel RN sowie das Schema K erweitert. Im folgenden verstärken wir schrittweise das System **K** bis zur Logik **S5** und kommen dann (pp. p. 127ff.) auf die Vollständigkeit von **S5** zurück.

5. Relationale Modelle und das modale System K

Wir haben schon eingangs darauf hingewiesen, daß der Begriff der Notwendigkeit auf sehr verschiedene Weise bestimmt werden kann und das zumindest einige dieser Bestimmungen auch davon abhängig sein können, wie unsere Welt beschaffen ist. Hier sind vier Beispiele indexikalischer Notwendigkeit:

- *Physikalische Notwendigkeit.* Die fundamentalen physikalischen Parameter (wie Gravitation oder Lichtgeschwindigkeit) haben in unserer Welt w bestimmte Werte. Wenn wir sagen, eine Aussage A gelte mit physikalischer Notwendigkeit, dann behaupten wir nicht, daß sie in allen möglichen Welten (im gesamten Bereich W) wahr sei, sondern in all denen, in denen die fundamentalen physikalischen Parameter und Gesetze wie in a sind. Wir quantifizieren also über einen Teilbereich $M(a) \subseteq W$, der abhängig von a bestimmt ist.
- *Epistemische Notwendigkeit.* Wir wissen vieles, jedoch nicht alles. Wenn Hans weiß, daß Peter in Frankfurt zuhause ist (P), dann ist das so (denn was nicht wahr ist, kann man nicht wissen), und dann blendet Hans alle Möglichkeiten aus, in denen es nicht so ist. Das heißt, $\neg P$ stellt für Hans keine ernsthafte Möglichkeit dar, oder, anders gesagt, in

allen ernsthaften Möglichkeiten, die Hans in Betracht zieht, ist P wahr. Statt von "ernsthaften Möglichkeiten" sprechen wir auch von *epistemischen Möglichkeiten.* Da für Hans die Aussage P in *allen* epistemischen Möglichkeiten ("Welten") wahr ist, ist P in diesem Sinne für Hans eine epistemische Notwendigkeit. In anderen als der unseren Welt weiß Hans andere Dinge. Was für Hans eine epistemische Möglichkeit in der einen Welt ist, muß keine solche in einer anderen Welt sein. Auch hier quantifizieren wir also über einen Teilbereich $M(a)$ aller Möglichkeiten W, der abhängig von einem Ausgangspunkt a bestimmt ist.

· *Deontische Notwendigkeit.* In der deontischen Modallogik interpretieren wir $\Box A$ als: A soll der Fall sein. Normen teilen den Bereich möglicher Welten in solche, die den Normen gehorchen (deontisch "ideale" Welten) und solche, die das nicht tun. Viele Normen sind konditional: Im Fall A soll B geschehen, im Fall C ist D geboten, und anderenfalls soll E sein. Andere sind in einem bestimmten Sinne indexikalisch: Daß alle glücklich sein sollen, bedeutet in einer Welt mit 4 Milliarden Menschen (1974) wohl etwas anderes als in einer Welt mit 7 Milliarden (2011). Wie also einem System von Normen zufolge eine Welt aussehen darf, wird in der Regel auch davon abhängig sein, wie sie tatsächlich ist. So ist zu erwarten, daß die Welten $M(a)$, die von a aus gesehen zulässig sind, verschieden sind von denen, die unter den Gegebenheiten einer anderen Welt, a', zulässig sind.

· *Temporale Notwendigkeit.* Schließlich ist unschwer zu erkennen, daß auch die Zeitoperatoren des letzten Kapitels sich in dieses Schema einfügen. Von einem Zeitpunkt a aus betrachtet, war etwas in der Vergangenheit notwendig (H), wenn es zu allen Zeitpunkten $M(a) = \{x \in W : x \prec a\}$ so war; die spiegelbildliche Formulierung gilt ebenso für die Zukunftsnotwendigkeit G.

Vielleicht ist *logische Notwendigkeit* so etwas wie absolute Notwendigkeit: Wahrheit in schlichtweg *allen* möglichen Welten. Dann wäre die Wahrheitsbedingung

$$(\dagger) \qquad a \models \Box A \text{ gdw } \forall x : x \models A$$

genau richtig. Andere wichtige Notwendigkeitsbegriffe sind aber nicht absolut, sondern im gerade illustrierten Sinne relativ. Für diese erwarten wir – wie in den Beispielen illustriert – eine Wahrheitsbedingung dieser Art:

$$(\ddagger) \qquad a \models \Box A \text{ gdw } \forall x : x \in M(a) \Rightarrow x \models A.$$

Nun kann es nicht Aufgabe des Logikers sein, die jeweiligen Mengen $M(a)$ zu bestimmen. (Das ist Aufgabe des Physikers, Metaphysikers oder Gesetzgebers, u.s.w.) Wohl aber kann er danach fragen, welche Strukturmerkmale eine Einschränkung der Menge der Welten – gleichgültig, wie diese im Detail ausfallen mag – haben kann und wie diese sich auf die logischen Eigenschaften des Notwendigkeitsbegriffs auswirken. In der Modallogik ist es üblich, solche Strukturmerkmale als Eigenschaften von Relationen anzugeben. Deshalb schreiben wir ab nun "Rax" für "$x \in M(a)$" und reformulieren (†) so:

$$(\Box) \qquad a \models \Box A \text{ gdw } \forall x : Rax \Rightarrow x \models A.$$

Jetzt können wir allgemeine Eigenschaften der Relation R in den Blick nehmen, wie Reflexivität, Symmetrie, Transitivität u.a.[5] Die Bedingung ($\Box$) enthält übrigens die "absolute" Bedingung (†) als einen Grenzfall: (†) resultiert aus ($\Box$) in dem Fall, in dem die Relation R universal in W ist, d.h. alle Punkte in W zueinander in Beziehung setzt.

Mehrere Begriffe von Notwendigkeit können offenbar gleichzeitig in einer Theorie behandelt werden. So könnte man Operatoren $\Box_\alpha$ und $\Box_\phi$ mit Hilfe zweier Relationen, R_α und R_ϕ, interpretieren. Wenn z.B. $\Box_\alpha$ arithmetische und $\Box_\phi$ physikalische Notwendigkeit repräsentieren, dann werden wir eine gewisse Interaktion zwischen den beiden Operatoren erwarten, beispielsweise

$$\Box_\alpha A \to \Box_\phi A$$

– alle mathematischen Notwendigkeiten sind auch physikalisch notwendig. Dieses Axiom wird gültig durch die Bedingung $R_\phi ab \Rightarrow R_\alpha ab$ (d.h. $R_\phi \subseteq R_\alpha$). Die Bedingung besagt, daß nur arithmetisch mögliche Alternativen zu a auch physikalisch mögliche Alternativen sind.

Wenn wir $\Box A$ deontisch interpretieren, d.h. im Sinne von "A *sollte* der Fall sein", dann setzt die entprechende Relation R_δ eine Welt a in Beziehung zu ihren "ethischen Vervollkommnungen", d.h. Welten, in denen es sich so verhält, wie es in a sein sollte. Daß etwas in einer Welt geboten ist, sagt dann so viel wie, daß es in allen "guten" Welten so ist. Nun sollten Normen nichts physikalisch Unmögliches verlangen:

$$\Box_\delta A \to \Diamond_\phi A.$$

Dieses Schema geht offenbar einher mit der Bedingung, daß unter den guten Alternativen zu einer Welt mindestens eine sein muß, die physikalisch realisierbar ist; kurz: $\forall a \exists b : R_\delta ab \;\&\; R_\phi ab$. Logiken über Sprachen, in denen verschiedene Modaloperatoren miteinander kombiniert werden können,

[5] Das geht natürlich im Prinzip auch mit der M-Notation in (‡). Dann wird Reflexivität, Raa, zu $a \in M(a)$ und Symmetrie, $Rab \Rightarrow Rba$, zu $b \in M(a) \Rightarrow a \in M(b)$ (jeweils für alle $a, b \in W$).

nennt man *multimodale* Logiken. Einfache Beispiele solcher Logiken haben wir im Kapitel II über Zeitlogik kennengelernt. In diesem Kapitel werden wir nicht weiter auf multimodale Logiken eingehen.

Unsere *Modelle* sehen jetzt so aus:

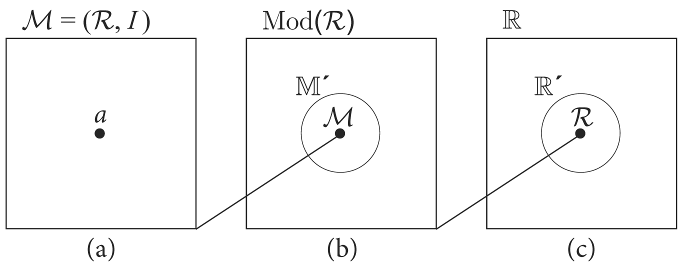
$$(W, R, I),$$

mit $R \subseteq W \times W$. Die Bedingungen für die auf I basierende Wahrmacherrelation $\models$ sind wie zuvor – nun erweitert um die relationale Bedingung ($\Box$).

- Die Modellkomponente (W, R) nennen wir einen *Kripke-Rahmen* und bezeichnen die Klasse aller Kripke-Rahmen mit $\mathbb{K}$.[6]
- Die *Theorie* der Kripke-Rahmen, $Th(\mathbb{K})$, ist diejenige Teilmenge der Formeln von $\mathcal{L}^{\Box}$, die wahr sind an allen Punkten in einem beliebigen Modell auf einem beliebigen Kripke-Rahmen.

Notation: Wenn $\mathbb{R}$ eine Klasse von Rahmen ist, dann sei $Th(\mathbb{R}) = \{A \in \text{FML} : \mathbb{R} \models A\}$. Statt $Th(\{\mathcal{R}\})$ schreiben wir $Th(\mathcal{R})$. Ähnlich notieren die Theorien von Modellen $\mathcal{M}$ und Modellklassen $\mathbb{M}$ mit $Th(\mathcal{M})$ bzw. $Th(\mathbb{M})$.

Die Definition von $Th(\mathbb{K})$ enthält eine dreifache Quantifikation. Sie ist dennoch leicht zu fassen, da es sich um eine Abfolge nur von All-Quantifikationen handelt: alle Punkte – alle Modelle – alle Rahmen. Für das bessere Verständnis des Folgenden bietet sich hier eine Gelegenheit, sich das zugrundeliegende Bild deutlich vor Augen zu stellen.



Das Bild veranschaulicht so etwas wie "Allgemeinheitsgrade der Wahrheit". In (a) sehen wir ein bestimmtes Modell $\mathcal{M}$ auf einem bestimmten Rahmen $\mathcal{R}$ mit einem Punkt a. Wir können nach der Wahrheit einer Formel A an diesem Punkt fragen oder auch nach der Wahrheit von A im Modell – dann

[6] Das ist die heute übliche Benennung. Warum damit kein Prioritätsanspruch in der Sache angedeutet sein kann, haben wir schon zuvor erklärt (pp. 81f.).

fragen wir, ob A an *allen* Punkten in $\mathcal{M}$ wahr ist. Nun ist $\mathcal{M}$ ein Modell unter vielen auf dem Rahmen $\mathcal{R}$. In (b) sehen wir $\mathcal{M}$ dargestellt als Element • einer Klasse $\mathbb{M}'$ von Modellen, die alle zur Klasse Mod($\mathcal{R}$) der Modelle auf $\mathcal{R}$ gehören. Auf dieser Ebene der Allgemeinheit können wir fragen, ob A in allen Modellen in $\mathbb{M}'$ oder sogar in allen Modellen auf $\mathcal{R}$, d.h. in Mod($\mathcal{R}$) wahr ist. Im letzten Fall sagen wir, daß A im Rahmen $\mathcal{R}$ wahr ist, denn es kommt nun offenbar auf die Wahl der Modelle nicht mehr an. In (c) erscheint $\mathcal{R}$ nun als zu einer Klasse $\mathbb{R}'$ von Rahmen gehörig und die Frage entsteht, ob alle Rahmen in dieser Klasse die Formel A wahr machen. Schließlich – der höchste Grad der Allgemeinheit – können wir noch fragen ob Rahmen überhaupt derart sind, daß sie A wahr machen, und umgekehrt: Welche Formeln sind in beliebigen Rahmen, d.h. in $\mathbb{R}$ wahr? Diese Grade der Wahrheit sind in der folgenden Definition zusammengefaßt.

Definition 5. Es sei $\mathbb{R}$ eine Klasse von Kripke-Rahmen und $\mathcal{M} = (W, R, I)$ sei ein Modell auf einem Rahmen (W, R) aus $\mathbb{R}$.

1. A ist *wahr an einem Punkt* $a \in W$ gdw $a \models A$ in $\mathcal{M}$.
2. A ist *wahr in einem Modell* $\mathcal{M}$ gdw A an allen Punkten in $\mathcal{M}$ wahr ist. Notation: $\mathcal{M} \models A$.
3. A ist *gültig in einem Rahmen* $\mathcal{R}$ gdw A in allen Modellen auf $\mathcal{R}$ wahr ist. Notation: $\mathcal{R} \models A$.
4. A ist *gültig in einer Klasse* $\mathbb{R}$ *von Rahmen* gdw A in allen $\mathcal{R} \in \mathbb{R}$, d.h. in allen Modellen auf beliebigen Rahmen in $\mathbb{R}$ wahr ist. Notation: $\mathbb{R} \models A$.
5. $Th(\mathbb{R}) = \{A : \mathbb{R} \models A\}$ ist die *Theorie* der Rahmenklasse $\mathbb{R}$.

Offensichtlich kann eine Formel wahr an einem Punkt sein, ohne wahr im Modell zu sein; sie kann wahr in einem Modell sein, ohne wahr in einem Rahmen zu sein, u.s.w. Vielleicht etwas weniger offensichtlich aber nicht minder wichtig ist in dieser Hinsicht der Status von Regeln. So ist jeder Punkt eines Modells unter Modus Ponens abgeschlossen; also ist auch Wahrheit im Modell, in einem Rahmen, und Gültigkeit in einer Klasse von Rahmen unter MP abgeschlossen. Die Punkte eines Modells sind jedoch nicht unter RN abgeschlossen; wohl aber ist Wahrheit in einem Modell u.s.w. so abgeschlossen.[7] Ferner ist der Begriff der Wahrheit in einem Modell nicht unter gleichförmiger Einsetzung abgeschlossen. Dazu genügt es, ein Modell $\mathcal{M}$ zu betrachten mit nur einem Punkt a so, daß $I(P, a) = 1$ und $I(Q, a) = 0$. In diesem Modell ist P wahr, d.h. $\mathcal{M} \models P$. Die in $\mathcal{M}$ falsche Formel Q resultiert aus P durch den einfachsten Fall einer gleichförmigen Einsetzung.

[7] Manchmal wird das so beschrieben: RN gilt nicht für "lokale", sondern nur für "globale" Wahrheit in einem Modell. Siehe dazu auch den nächsten Abschnitt über den Status des Deduktionstheorems.

Wenn wir jedoch übergehen zu Wahrheit in einem Rahmen (und Gültigkeit in einer Klasse von Rahmen), dann werden die Verschiedenheiten der Interpretationen wegquantifiziert und die Regel der gleichförmigen Einsetzung überträgt den Wahrheits- bzw. Gültigkeitsstatus von der Prämisse an die Konklusion.

Die naheliegendste Frage überhaupt ist die Frage nach Wahrheit im allgemeinsten Sinne, d.h. nach der Theorie der Klasse $\mathbb{K}$ *aller* Kripke-Rahmen. Wir fragen hier nach der Logik von Notwendigkeit in einem nicht weiter spezifizierten Sinne. Wie läßt sich diese Theorie axiomatisch beschreiben?

Die kleinste normale modale Logik K. Wir haben oben (p. 80) schon beschrieben, was wir unter einer *Modallogik* verstehen wollen. Das ist jede Erweiterung **L** der klassischen Aussagenlogik **KL** in einer modalen Sprache $\mathcal{L}^\Box$, welche unter Modus Ponens und gleichförmiger Einsetzung abgeschlossen ist.[8] Eine Modallogik nennen wir *normal* (vgl. p. 90), wenn sie darüberhinaus unter der Normalitätsregel RK abgeschlossen ist:

RK $$\frac{(A_1 \wedge \cdots \wedge A_n) \to A}{(\Box A_1 \wedge \cdots \wedge \Box A_n) \to \Box A} \quad (n \geq 0).$$

Daß RK aus K und RN folgt, wurde oben gezeigt. In umgekehrter Richtung ist RN der $(n = 0)$-Fall von RK und der Schluß von $(A \to B) \wedge A \to B$ auf $\Box(A \to B) \wedge \Box A \to \Box B$ ist ein $(n{=}2)$-Fall von RK, aus dessen Konklusion das Schema K schon in **KL** durch Exportation für $\to$ folgt. Da RK also äquivalent ist zur Kombination von RN mit K, läßt sich die kleinste normale Modallogik auch so beschreiben:

Das System **K**

τ	Jede Tautologie
K	$\Box(A \to B) \to (\Box A \to \Box B)$
Regeln	$\dfrac{A \quad A \to B}{B}\text{MP} \qquad \dfrac{A}{\Box A}\text{RN}$

Da wir die Beobachtung später gut brauchen können, nutzen wir die Gelegenheit, um zu bemerken, daß die Regel RK (für Konjunktion und Pfeil) eine Entsprechung findet in einer Regel für Annahmenmengen und Ableitbarkeit.

[8] "Modale Sprache" deutet hier die Interpretation von $\Box$ als All-Quantor über die Trägermenge eines Kripke-Modells an.

Lemma 6. *In* **K** *gilt die folgende Regel für* $\vdash$:

RK$\vdash$
$$\frac{A_1, \ldots, A_n \vdash B}{\Box A_1, \ldots, \Box A_n \vdash \Box B}\ (n \geq 0).$$

Beweis. Sei $X = \{A_1, \ldots, A_n\}$ und $\Box X = \{\Box A_1, \ldots, \Box A_n\}$. Induktion über die Länge k einer Ableitung von B aus X, was hier so angedeutet sei:

$$X \vdash^k B.$$

Im Falle $k = 1$ muß B entweder (a) ein Axiom oder (b) in X sein. Wenn (a), dann $\vdash B$ und also nach RN, $\vdash \Box B$, woraus wir aufgrund der Monotonie von $\vdash$ auf $\Box X \vdash \Box B$ schließen dürfen. Wenn (b), dann fragen wir hier nach $\Box X \vdash \Box B$ mit $B \in X$, was aufgrund der Reflexivität (und Monotonie) von $\vdash$ gilt.

Im Induktionsschritt nehmen wir nun an, daß für alle Längen $i < k$ gilt:

IA.
$$X \vdash^i B \Rightarrow \Box X \vdash \Box B.$$

Wir nehmen ferner an, daß $X \vdash^k B$ (zz: $\Box X \vdash \Box B$). Dann ergeben sich die zwei Möglichkeit (a) und (b), wie gerade behandelt, sowie die weiteren Fälle, daß (c) B durch MP oder (d) B durch RN gewonnen wurde.

Der Fall (c), MP, stellt sich so dar: In die Position k der Ableitung durften wir B eintragen, da es Positionen x und y vor k und eine Formel B gibt mit

$$X \vdash^y C \to B \quad \text{sowie} \quad X \vdash^x C.$$

Darauf können wir die IA anwenden und erhalten

$$(*)\ \Box X \vdash \Box(C \to B) \quad \text{und} \quad (\dagger)\ \Box X \vdash \Box C.$$

Nun folgt aus $(*)$ aufgrund von K (und MP), daß

$$\Box X \vdash \Box C \to \Box B,$$

woraus aufgrund von $(\dagger)$ (und MP) folgt, daß $\Box X \vdash \Box B$, wie gewünscht.

Im Fall (d), RN, ist B von der Form $\Box C$ und es gibt ein $x < k$ mit $X \vdash^x C$. Darauf wenden wir die IA an und erhalten $\Box X \vdash \Box C$. Eine Ableitung von $\Box C$ können wir mittels RN zu einer Ableitung von $\Box\Box C$ ($=\Box B$) verlängern. So erhalten wir $\Box X \vdash \Box B$, wie gewünscht. ∎

Anmerkung. Man beachte, daß wir uns zu RK$\vdash$ nicht auf folgende einfache Weise verhelfen konnten: Aus $A_1, ..., A_n \vdash B$ gewinnen wir durch fortgesetzte Anwendung des Deduktionstheorems und Importation zunächst $\vdash A_1 \wedge \cdots \wedge A_n \rightarrow B$. Darauf wenden wir RK an und bringen dann die Konjunkte wieder auf die rechte Seite von $\vdash$. Wie schon bemerkt, ist das Deduktionstheorem für **K** nicht zulässig!

LEMMA 7. RK$\vdash$ *ist äquivalent zu* RN *und*

RR $\vdash$
$$\frac{A, B \vdash C}{\Box A, \Box B \vdash \Box C}.$$

BEWEIS. Die Richtung RK$\vdash$ $\Rightarrow$ RR$\vdash$ & RN ist trivial. Für die umgekehrte Richtung induzieren wir über die Anzahl $n \geq 0$ der Annahmen.

Im Fall $n = 0$ haben wir $\vdash B$ und wenden RN zur gewünschten Konklusion $\vdash \Box B$ an.

Unsere IA ist für $m < n$,

IA.
$$\frac{A_1, \ldots, A_m \vdash B}{\Box A_1, \ldots, \Box A_m \vdash \Box B}.$$

Wir nehmen an, daß (1) $A_1, \ldots, A_m, A_{m+1} \vdash B$. Daraus folgt schon für **KL**

(2)
$$A_1, \ldots, A_m \wedge A_{m+1} \vdash B.$$

Im Schritt von (1) nach (2) haben wir die Anzahl der Annahmen auf m reduziert, weshalb wir von der IA Gebrauch machen dürfen:

(3)
$$\Box A_1, \ldots, \Box A_{m-1}, \Box(A_m \wedge A_{m+1}) \vdash \Box B.$$

Nun haben wir aufgrund von RR$\vdash$,

(4)
$$\Box A_m, \Box A_{m+1} \vdash \Box(A_m \wedge A_{m+1}).$$

(Denn ganz allgemein können wir nach RR$\vdash$ von $A, B \vdash A \wedge B$ übergehen zu $\Box A, \Box B \vdash \Box(A \wedge B)$.) So folgt aus (3) und (4) durch Schnitt

$$\Box A_1, \ldots, \Box A_m, \Box A_{m+1} \vdash \Box B,$$

wie gewünscht. ∎

Eine alternative Axiomatisierung von **K** ergibt sich aus dem folgenden

Lemma 8. *Eine Modallogik ist genau dann normal, wenn sie alle Tautologien enthält (τ), unter der Regel*

RE. $$\frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B}$$

abgeschlossen ist, sowie alle Instanzen dieser Schemata enthält:

N. $$\Box\top$$

C. $$\Box A \wedge \Box B \to \Box(A \wedge B)$$

M. $$\Box(A \wedge B) \to \Box A \wedge \Box B$$

Beweis. Von links nach rechts folgt die Beobachtung recht unmittelbar durch Anwendung von RK mit $n = 1$ bzw. $n = 2$. In der umgekehrten Richtung leiten wir RN und K ab.

Ad RN: Wenn $\vdash A$, dann $\vdash A \leftrightarrow \top$. Also nach RE, $\vdash \Box A \leftrightarrow \Box\top$. Da $\vdash \Box\top$, so folgt $\vdash \Box A$.

Ad K: Wir beweisen zunächst den Abschluß unter der Regel

RM. $$\frac{A \to B}{\Box A \to \Box B}$$

(1)	$A \to B$	Annahme
(2)	$\Box(A \to B)$	1, RN
(3)	$\Box A \to \Box(A \to B) \wedge \Box A$	2, τ etc.
(4)	$\Box(A \to B) \wedge \Box A \to \Box((A \to B) \wedge A)$	C
(5)	$\Box((A \to B) \wedge A) \to \Box(A \wedge B)$	τ, RE
(6)	$\Box A \to \Box(A \wedge B)$	3,4,5
(7)	$\Box(A \wedge B) \to \Box B$	M, τ etc.
(8)	$\Box A \to \Box B$	6,7

Sodann:

(1)	$(A \to B) \wedge A \to B$	τ
(2)	$\Box((A \to B) \wedge A) \to \Box B$	1, RM
(3)	$\Box(A \to B) \wedge \Box A \to \Box B$	2, C, τ etc.
(4)	$\Box(A \to B) \to (\Box A \to B)$	3, τ etc

■

Es stellt sich heraus, daß alle Modelle auf Kripke-Rahmen die Eigenschaft haben, neben den klassischen Tautologien das Axiom K zu verifizieren und daß die Menge der in einem solchen Modell verifizierten Formeln unter Modus Ponens und RN abgeschlossen ist. Da das Axiom und die Regeln hier im schematischen Sinne verstanden sind, folgt unmittelbar auch die Eigenschaft, unter gleichförmiger Einsetzung abgeschlossen zu sein. Das ist im wesentlichen der Gehalt des folgenden Satzes. Zusammen mit dessen Umkehrung kommen wir so zu dem Resultat, daß das System **K** genau die Theorie der Kripke-Rahmen im allgemeinsten Sinne beschreibt, also $Th(\mathbb{K}) = \mathbf{K}$.

SATZ 9. (RICHTIGKEIT) *Das System* **K** *ist korrekt für die Klasse* $\mathbb{K}$ *der Kripke-Rahmen: Wenn* $\vdash A$ *in* **K**, *dann* $A \in Th(\mathbb{K})$, *d.h.* $\mathcal{M} \models A$ *in allen Modellen* $\mathcal{M}$ *auf beliebigen Kripke-Rahmen.*

BEWEIS. Es genügt, den Beweis von Satz 2 um die Verifikation des Schemas K jetzt unter der *relationalen* Bedingung ($\Box$) zu ergänzen:

1	(1)	$a \models \Box(A \to B)$	Annahme
2	(2)	$a \models \Box A$	Annahme
3	(3)	Rab	Annahme/zz $b \models B$
1	(4)	$\forall x : Rax \Rightarrow x \models A \to B$	1 ($\Box$)
1	(5)	$\forall x : Rax \;\&\; x \models A \Rightarrow x \models B$	4 ($\to$)
2	(6)	$\forall y : Ray \Rightarrow y \models A$	2 ($\Box$)
2,3	(7)	$b \models A$	3,6
1,2,3	(8)	$b \models B$	3,5,7

■

Es folgt die Umkehrung des Satzes:

SATZ 10. (VOLLSTÄNDIGKEIT) *Das System* **K** *ist vollständig für die Klasse* $\mathbb{K}$ *der Kripke-Rahmen: Wenn* $A \in Th(\mathbb{K})$, *dann* $\vdash A$ *in* **K**.

Um den Satz zu beweisen, kontraponieren wir und nehmen an, daß eine Formel A *kein* Theorem von **K** sei. Dazu müssen wir ein Modell auf einem Rahmen in $\mathbb{K}$ finden, welches A an einem Punkt falsch macht. Man beachte, daß es für den Satz genügen würde, wenn es für *jedes* Nicht-Theorem (mindestens) *ein* falsifizierendes Modell gäbe. Wir werden etwas Stärkeres zeigen: Es gibt *ein* Modell, welches *jedes* Nicht-Theorem falsifiziert. Dieses Modell werden wir das *kanonische Modell* $\mathcal{M}^{\mathbf{K}}$ für **K** nennen. Da dieses Modell umgekehrt auch jedes Theorem von **K** wahr macht, stellt es **K** richtig und vollständig dar: Es macht *genau* die Theoreme von **K** wahr und ist in diesem Sinne ein *charakteristisches Modell* für **K**.

Wie sieht nun das kanonische Modell für **K**,

$$\mathcal{M}^{\mathbf{K}} = (W^{\mathbf{K}}, R^{\mathbf{K}}, I^{\mathbf{K}}),$$

aus? Die Punkte in $W^{\mathbf{K}}$ sind Formelmengen. Genauer: $W^{\mathbf{K}}$ ist der Bereich aller Formelmengen, die im Sinne von **K** maximal konsistent sind.

- Eine Menge $X \subseteq \mathrm{FML}$ ist genau dann *maximal konsistent* im Hinblick auf eine Logik **L**, wenn

M1. $X \nvdash_{\mathbf{L}} \perp$ (Konsistenz), und
M2. $\forall A$: wenn $A \notin X$, dann $X, A \vdash_{\mathbf{L}} \perp$ (Maximalität).

Wenn wir in diesem Abschnitt von maximaler Konsistenz sprechen, dann ist diese immer bezüglich der Logik **K** gemeint.

Maximal konsistente Mengen X haben einige Eigenschaften, die den Schlüssel zum Vollständigkeitssatz abgeben. Dazu gehören diese:

M3. $\neg A \in X$ gdw $A \notin X$.
In der LR-Richtung drückt M3 noch einmal die Konsistenz M1 von X aus. In umgekehrter Richtung nehme man an, daß sowohl $A \notin X$ als auch $\neg A \notin X$. Dann $X, A \vdash \perp$ und $X, \neg A \vdash \perp$, woraus $X \vdash \perp$ folgt – entgegen M1!

M4. $A \in X$ gdw $X \vdash A$.
Die LR-Richtung ist trivial. Für RL nehmen wir an, daß $X \vdash A$ und $A \notin X$, d.h. $X, A \vdash \perp$. Dann erhalten wir durch Schnitt $X \vdash \perp$ – Widerspruch zu M1!

M5. $A \wedge B \in X$ gdw $A \in X$ und $B \in X$.
Wenn $A \wedge B \in X$, dann $X \vdash A \wedge B$, woraus durch $\wedge$-Beseitigung folgt, daß $X \vdash A$ und also (nach M4), daß $A \in X$; ebenso für B. Für die umgekehrte Richtung kehren wir das Argument um und benutzen $\wedge$-Einführung.

Es sei nun $W^{\mathbf{K}}$ die Menge aller maximal konsistenten Mengen. Dann läßt sich gleich eine weitere Beobachtung über den Schnitt aller Mengen in $W^{\mathbf{K}}$ anfügen:

M6. $\bigcap W^{\mathbf{K}} = \mathbf{K}$.
Die $\supseteq$-Richtung ist einfach: Es sei X eine beliebige Menge in $W^{\mathbf{K}}$ und A sei in **K**, also $\vdash A$. Dann $X \vdash A$ (Monotonie von $\vdash$!) und also $A \in X$ (M4!). — Für die umgekehrte, d.h. die $\subseteq$-Richtung brauchen wir das folgende Lemma.

LEMMA 11. (Lindenbaum) *Jede* (**L**-) *konsistente Formelmenge X läßt sich zu einer maximal konsistenten Menge (bezüglich* **L***) erweitern.*

Den Beweis haben wir in der Einleitung gegeben für eine ganze Klasse von Logiken **L**, zu denen insbesondere jede Erweiterung der klassischen Aussagenlogik **KL**, also auch **K**, gehört. ∎

Nun fahren wir fort mit der Verifikation von M6, d.h. wir zeigen, daß $\bigcap W^{\mathbf{K}} \subseteq \mathbf{K}$, oder, in anderen Worten:

$$\text{wenn } \forall X \in W^{\mathbf{K}} : A \in X, \text{ dann } A \in \mathbf{K}.$$

Wir beweisen die Kontraposition. Angenommen eine Formel A (eine bestimmte Formel, kein Schema!) ist kein Theorem von $\mathbf{K}$. Dann wissen wir, daß (in $\mathbf{K}$) $\{\neg A\} \nvdash \bot$, d.h. $\{\neg A\}$ ist eine konsistente Formelmenge (bezüglich $\mathbf{K}$). Nach Lemma 11 läßt sich $\{\neg A\}$ zu einer Menge X in $W^{\mathbf{K}}$ erweitern. Diese Menge X enthält $\neg A$ und daher (M3!) $A \notin X$. So haben wir, wie gefordert, gezeigt, daß es unter der Annahme $A \notin \mathbf{K}$ eine Menge X in $W^{\mathbf{K}}$ gibt mit $A \notin X$.

Für die abschließende Beobachtung über maximal konsistente Mengen wollen wir eine abkürzende Notation vereinbaren. Wenn in einer Formelmenge X $\Box$-Formeln vorkommen, z.B. $\Box A$, dann können wir in einer anderen Menge $X^{\Box}$ alle Formeln einsammeln, von denen in X behauptet wird, sie seien notwendig, also z.B. A. Die Menge $X^{\Box}$ ist gewissermaßen die "Entnotwendigung" von X:

$$X^{\Box} = \{A : \Box A \in X\}.$$

Wenn nun X eine maximal konsistente Menge ist, dann können wir folgendes beobachten:

M7. $\Box A \in X$ gdw für alle maximal konsistenten Y: $X^{\Box} \subseteq Y \Rightarrow A \in Y$.

Warum es wert ist, die Beobachtung M7 festzuhalten, mag sich nicht unmittelbar erschließen. Deshalb sei hier eine Lesart angegeben, die andeutet, wohin die Reise gleich gehen wird.

(M7.) In einer Menge X in $W^{\mathbf{K}}$ ist eine Formel $\Box A$ "wahr" ($\Box A \in X$) gdw A in allen Mengen Y in $W^{\mathbf{K}}$ "wahr" ist, die zu X in dieser *Relation* stehen: $X^{\Box} \subseteq Y$. (Das sollte einen an die Wahrheitsbedingung für $\Box$-Formeln denken lassen!)

Es folgt der Beweis von M7:

Die LR-Richtung folgt unmittelbar aus der Definition von $X^{\Box}$: Wenn $\Box A \in X$, dann ist A in $X^{\Box}$ und also auch in jeder Obermenge Y von $X^{\Box}$. — Für die umgekehrte Richtung nehmen wir an, daß $\Box A \notin X$. (Wir zeigen dann die Verneinung der rechten Seite.) Aus der Annahme folgt, daß $A \notin X^{\Box}$. Es sei nun $X' = X^{\Box} \cup \{\neg A\}$. Diese Menge ist konsistent. (Angenommen, das sei nicht so, d.h. $X^{\Box}, \neg A \vdash \bot$, also $X^{\Box} \vdash A$. Dann gibt es $A_1, \ldots, A_n$ in $X^{\Box}$ so, daß $A_1, \ldots, A_n \vdash A$. Darauf wenden wir RK$\vdash$

an und erhalten $\Box A_1, \ldots, \Box A_n \vdash \Box A$. Da aber jedes $A_i \in X^{\Box}$, so ist jedes $\Box A_i \in X$. Daher $X \vdash \Box A$ und so (M4!) $\Box A \in X$ – im Widerspruch zu unserer Voraussetzung.) Die konsistente Menge X' können wir zu einer maximal konsistenten Menge Y erweitern (Lemma 11). Da $\neg A \in Y$, so ist (M3!) $A \notin Y$. Wir sind also, wie gewünscht, von der Annahme $\Box A \notin X$ zur Verneinung der rechten Seite gekommen: $\exists Y$ so, daß Y maximal konsistent ist, $X^{\Box}$ ist in Y enthalten, und $A \notin Y$.

Die Menge $W^{\mathbf{K}}$ des kanonischen Modells für **K** haben wir bestimmt und einige wichtige Eigenschaften ihrer Elemente genannt. Die kanonische Relation $R^{\mathbf{K}}$ sei nun so wie in (M7) angedeutet, definiert: Für alle $X, Y \in W^{\mathbf{K}}$,

$$(R^{\mathbf{K}}) \qquad R^{\mathbf{K}}XY \text{ gdw } X^{\Box} \subseteq Y.$$

Es ist klar, daß $(W^{\mathbf{K}}, R^{\mathbf{K}})$ ein Kripke-Rahmen ist. Denn $W^{\mathbf{K}}$ ist eine nicht-leere Menge und $R^{\mathbf{K}}$ ist eine binäre Relation in $W^{\mathbf{K}}$. Wir ergänzen diesen Rahmen zu einem Kripke-Modell, in dem wir eine Interpretation $I^{\mathbf{K}} : \mathrm{ATM} \times W^{\mathbf{K}} \longrightarrow \{0, 1\}$ definieren:

$$(I^{\mathbf{K}}) \qquad I^{\mathbf{K}}(P, X) = 1 \text{ gdw } P \in X,$$

für beliebige Atome P. Es folgt der Hauptsatz für das kanonische Modell:

Satz 12. *Im kanonischen $\mathcal{M}^{\mathbf{K}} = (W^{\mathbf{K}}, R^{\mathbf{K}}, I^{\mathbf{K}})$ für* **K** *gilt für alle Formeln A und $X \in W^{\mathbf{K}}$: $X \models A$ gdw $A \in X$.*

Beweis. Induktion über den Aufbau von A. Die Basis $A = P$ ist durch die Definition $(I^{\mathbf{K}})$ gegeben. Für die Fälle $A = \neg B$ und $A = B \wedge C$ greifen wir auf M3 bzw. M5 zurück. Im Fall $A = \Box B$ ist zu zeigen, daß

$$X \models \Box B \text{ gdw } \forall Y : R^{\mathbf{K}}XY \Rightarrow X \models B.$$

Die Induktionshypothese dürfen wir auf $X \models B$ anwenden und $R^{\mathbf{K}}XY$ lösen wir nach Definition $(R^{\mathbf{K}})$ auf. So erhalten wir

$$X \models \Box B \text{ gdw } \forall Y : X^{\Box} \subseteq Y \Rightarrow B \in X.$$

Die rechte Seite bedeutet aber nach M7 nichts anderes als $\Box B \in X$. ∎

Korollar 13. *Das kanonische Modell $\mathcal{M}^{\mathbf{K}}$ ist charakteristisch für* **K**.

Beweis. Sei $Th(\mathcal{M}^{\mathbf{K}}) = \{A : \mathcal{M}^{\mathbf{K}} \models A\}$. Nun gilt $\mathcal{M}^{\mathbf{K}} \models K$ gdw $\forall X \in W^{\mathbf{K}}, X \models A$. Nach dem Satz ist die rechte Seite genau dann wahr, wenn $\forall X \in W^{\mathbf{K}}, A \in X$, d.h. $A \in \bigcap W^{\mathbf{K}}$. Wir wissen (M6!) daß $\bigcap W^{\mathbf{K}} = \mathbf{K}$. Also ist $Th(\mathcal{M}^{\mathbf{K}}) = \mathbf{K}$. ∎

Der Vollständigkeitssatz 10 folgt nun unmittelbar aus diesem Korollar: Wenn A kein Theorem von **K** ist, dann gibt es ein Modell auf einem Kripke-Rahmen – nämlich das kanonische Modell $\mathcal{M}^{\mathbf{K}}$ –, in welchem die Formel A nicht wahr ist. Kontraponiert: Wenn es in der Klasse $\mathbb{K}$ der Kripke-Rahmen kein A falsifizierendes Modell gibt, A also in $\mathbb{K}$ gültig ist, dann ist A ein Theorem von **K**.

Wenn wir das Argument für die Vollständigkeit von **K** bezüglich der Klasse $\mathbb{K}$ der Kripke-Rahmen Revue passieren lassen, dann stellen wir fest, daß nur an wenigen Stellen etwas davon abhing, daß es hier um die *kleinste* normale Modallogik **K** und um die Klasse *aller* Kripke-Rahmen ging.

· Der Begriff einer maximal konsistenten Formelmenge wurde im Hinblick auf Ableitbarkeit $\vdash$ in beliebigen Logiken **L** definiert (M1 und M2).

· Die Eigenschaften M3–5 griffen zurück auf ganz allgemeine Eigenschaften der Relation $\vdash$.

· Für M6 haben wir das Lindenbaum-Lemma zwischengeschaltet. Auch dieses beruht auf allgemeinen Eigenschaft von $\vdash$.

· Allein in M7 ging es ein wenig modal zu. Hier mußten wir auf RN und K zurückgreifen. Das bedeutet, daß wir M7 für jede Modallogik, die über diese Ressourcen verfügt, also für jede normale Modallogik die Eigenschaft M7 nachweisen können.[9]

· Für den Beweis des Satzes 12 und sein Korollar wurde nur noch auf M1–7 zurückgegriffen.

Eigentlich haben wir also ein stärkeres Resultat bewiesen, nämlich

Satz 14. *Wenn* **L** *eine normale Modallogik ist, dann gilt im kanonischen Modell* $\mathcal{M}^{\mathbf{L}}$ *für alle Formeln* A *und* $X \in W^{\mathbf{K}}$*:* $X \models A$ *gdw* $A \in X$*.*

Daraus folgt als Korollar, daß $\mathcal{M}^{\mathbf{L}}$ charakteristisch für **L** ist.

Wenn nun die Logik **L** eine echte Erweiterung von **K** ist, dann kommen Eigenheiten dieser Logik erst ins Spiel, wenn wir für die Vollständigkeit

[9] Wir haben auch von der Regel RK$\vdash$ Gebrauch gemacht. Die Zulässigkeit dieser Regel für **K** haben wir durch eine Induktion über die Länge einer Ableitung bewiesen. Solche Induktionen sind nicht stabil gegen Erweiterungen um beliebige Regeln, denn neue Regeln zwingen zur Berücksichtigung neuer Ableitungsmöglichkeiten. Regeln, die, wenn **K** hinzugefügt, RK$\vdash$ unzulässig machen würden, sind aber nicht sehr natürlich (d.h. wohl nur mit dieser Absicht ersonnen) und spielen daher in den üblichen Logiken keine Rolle. Für alle Modallogiken, die wir hier betrachten, ist RK$\vdash$ eine gute Regel. Im Übrigen könnten wir für den Vollständigkeitsbeweis RK$\vdash$ umgehen, indem wir maximale Konsistenz etwas anders definieren; vgl. z.B. Chellas [53] oder Humberstone [144, §2.4]. Diese alternative Definition von Konsistenz resultiert aus einer alternativen Ableitbarkeitsbeziehung, die wir im nächsten Abschnitt besprechen werden.

von $\mathbf{L}$ im Hinblick auf eine irgendwie eingeschränkte Klasse $\mathbb{L}$ von Rahmen argumentieren. Hier ist noch einmal das schematische Argument:

- Angenommen $A \notin \mathbf{L}$. (Wir wollen beweisen, daß es dann ein Modell *auf Rahmen in* $\mathbb{L}$ gibt so, daß A in diesem Modell nicht wahr ist.)
- Wir bilden das kanonische Modell $\mathcal{M}^{\mathbf{L}}$.
- Wir wissen, daß $\mathcal{M}^{\mathbf{L}} \not\models A$.
- Aber ist $\mathcal{M}^{\mathbf{L}}$ ein Modell *auf einem Rahmen in* $\mathbb{L}$?

Hier brauchen wir einen Nachweis, daß $\mathcal{M}^{\mathbf{L}}$ die nötigen Bedingungen für Rahmen in $\mathbb{L}$ erfüllt – und dieser Nachweis wird sich auf solche Axiome von $\mathbf{L}$ stützen, die über $\mathbf{K}$ hinausgehen.

Damit haben wir eine Blaupause skizziert für Richtigkeits- und Vollständigkeitsnachweise für Erweiterungen von $\mathbf{K}$. Wir betrachten hier der Einfachheit halber den Fall, daß wir $\mathbf{K}$ um nur ein Axiom Ax erweitern.

- Wir suchen nach einer hinreichenden Bedingung Bed so, daß Ax gültig ist in der Klasse $Bed(\mathbb{K})$ der Rahmen, welche diese Bedingung erfüllen. Auf der Grundlage des Richtigkeitssatzes für $\mathbf{K}$ beweisen wir so die Richtigkeit von $\mathbf{K} + Ax$ für $Bed(\mathbb{K})$.
- Für die Vollständigkeit von $\mathbf{K} + Ax$ bzgl. $Bed(\mathbb{K})$ müssen wir nur zeigen, daß das kanonische Modell für $\mathbf{K} + Ax$ (genauer: die Rahmenkomponente des Modells) die Bedingung Bed erfüllt. (Der Rest geht wie für $\mathbf{K}$ vorgeführt.) Das wird nur gelingen, wenn Ax hinreichend für den Nachweis von Bed ist.
- Axiom und Bedingung müssen also genau zueinander passen: Die Bedingung muß hinreichend für das Axiom sein (Richtigkeit), und das Axiom muß hinreichend für die Bedingung sein (Vollständigkeit). Wenn beides zutrifft, dann *korrespondieren* Axiom und Bedingung.

Solche Korrespondenznachweise haben uns schon im Kapitel II.8 über Zeitlogik beschäftigt und wir werden ihnen weiter im Abschnitt 8 nachgehen. Zunächst jedoch kehren wir kurz zum Thema der Ableitbarkeit in der Modallogik zurück.

6. Gültige Schlüsse und das modale Deduktionstheorem

Bevor wir uns möglichen Erweiterungen von **K** zuwenden, wollen wir noch einmal die Frage nach einem geeigneten Deduktionstheorem für normale Modallogiken aufnehmen. Wir haben oben (p. 87) beobachtet, daß die Regel RN dazu führt, daß das Deduktionstheorem der **KL**,

(Ded) $$X, A \vdash B \;\Rightarrow\; X \vdash A \to B,$$

in der Modallogik nicht mehr gelten kann.[10] (Die Umkehrung von (Ded) gilt natürlich weiter uneingeschränkt.) Denn wenn wir eine kontingente Annahme A machen, dann erlaubt uns RN im nächsten Schritt einer Ableitung, $\Box A$ aufzuschreiben, mit dem Resultat $A \vdash \Box A$. Würden wir jetzt (Ded) anwenden, so erhielten wir $\vdash A \to \Box A$. Das ist aber sicher keine gültige Formel. (Das kleinste Gegenmodell hat zwei Punkte: $a \models A$, Rab und $b \not\models A$.) Also (Vollständigkeit!) ist $A \to \Box A$ kein Theorem von **K**.

Schon die Sequenz $A \vdash \Box A$ erregt vielleicht Verdacht. Kann man aus kontingenten Annahmen auf notwendige Wahrheiten schließen? Ist so etwas nicht ein elementares Beispiel eines *ungültigen* Schlusses? Nach dem bisherigen Stand unserer Definitionen können wir diese letzte Frage eigentlich noch gar nicht entscheiden. Denn wir haben in Definition 5 wohl definiert, was eine gültige *Formel* ist (in einer Klasse von Rahmen), nicht jedoch, was ein gültiger *Schluß* ist. Es ist an der Zeit, dies nachzuholen.

Gültige Schlüsse übertragen die Wahrheit der Prämissen auf die Konklusion. Sie tun dies mit einer gewissen Zwangsläufigkeit, d.h. "ohne Ausnahme". Diesen Aspekt der Zwangsläufigkeit holen wir ein durch eine Allquantifikation über einen geeigneten Bereich von Möglichkeiten, in denen die Prämissen und die Konklusion eines Schlusses einen Wahrheitswert tragen. Wir sagen: B folgt (gültig) aus A, wenn in jeder Möglichkeit für A wahr zu sein, auch B wahr ist. In der klassischen Aussagenlogik waren die Möglichkeiten Wahrheitswertverteilungen über die Atome einer Sprache. Solche Verteilungen sind die Modelle $\mathcal{M} = (\{0,1\}, I)$ einer wahrheitsfunktionalen Sprache. Gültige Folgerung $\models$ war demnach so definiert:

(*) $X \models A$ gdw $\forall \mathcal{M}$: wenn $\mathcal{M} \models X$, dann $\mathcal{M} \models A$.

Nun werden an den Punkten eines Kripke-Modells (W, R, I) Wahrheitswerte über die Atome genauso wie in Modellen der klassischen Aussagenlogik verteilt. In diesem Sinne ist ein Punkt $a \in W$ selbst "wie ein Modell". Wenn

[10] Vgl. den Aufsatz "Does the Deduction Theorem fail for modal logic?" [124], in dem der Frage sehr sorgfältig und detailreich nachgegangen wird. Die richtige Antwort ist natürlich ein klares "Jein", abhängig davon, was unter dem Deduktionstheorem verstanden wird und wie die Modallogik aussieht, für welche die Frage gestellt wird.

Folgerung allgemein Wahrheitsübertragung ist, dann liegt es jetzt nahe, Folgerung in einem Modell als "lokale" Wahrheitsübertragung zu definieren – die *loci* sind hier die Punkte des Modells:

(lok) $X \models^{\ell} A$ in $\mathcal{M}$ gdw $\forall a \in W$: wenn $a \models X$, dann $a \models A$.

Sodann quantifizieren wir die Eigenheiten einzelner Kripke-Modelle weg, indem wir Folgerung in *allen* Kripke-Modellen (auf einer Klasse $\mathbb{R}$ von Rahmen) in den Blick nehmen:

(Lok) $X \models^{\ell} A$ in $\mathbb{R}$ gdw $\forall \mathcal{M} \in \mathbb{M}(\mathbb{R}) : X \models^{\ell} A$ in $\mathcal{M}$.

Nach dieser Definition *lokaler Folgerung* müssen wir allerdings die Regel RN, d.h. den Übergang von A zu $\Box A$ als eine ungültige Folgerung beschreiben. Denn, wie schon anläßlich der Definition 5 beobachtet, ist Wahrheit an einem Punkt in einem Modell, also *lokale* Wahrheit, nicht unter der Regel RN abgeschlossen. Erst der Begriff der Wahrheit in einem Modell, d.h. Wahrheit an allen Punkten des Modells – *globale* Wahrheit, wie wir auch sagen können – ist so abgeschlossen. Vielleicht ist es also diese Art von Wahrheit, an deren Übertragung uns beim Folgern liegen sollte. In diesem Sinne folgt A aus X in einem Modell, wenn die globale Wahrheit von X, die globale Wahrheit von A impliziert, d.h.

(glob) $X \models^{g} A$ in $\mathbb{M}$ gdw: wenn $\mathcal{M} \models X$, dann $\mathcal{M} \models A$.

Auch hier wollen wir schließlich von den Eigenheiten einzelner Modelle absehen und erhalten so eine Definition *globaler Folgerung* in einer Klasse $\mathbb{R}$ von Rahmen:

(Glob) $X \models^{g} A$ (in $\mathbb{R}$) gdw $\forall \mathcal{M} \in \mathbb{M}(\mathbb{R})$: wenn $\mathcal{M} \models X$, dann $\mathcal{M} \models A$.

Daß auch diese Definition der Vorlage (*) auf natürliche Weise folgt, ist unschwer zu erkennen; sie folgt der Vorlage sogar in einem engeren Sinne als (Lok) dies tut. Denn obgleich die Punkte eines Kripke-Modells selbst "wie" Modelle sind, so sind sie es doch nicht im wörtlichen Sinne.

Die beiden Definitionen bestimmen die Relation der Folgerung auf sehr verschiedene Weise. Lokale Folgerung ist die stärkere Relation, d.h.

$$\models^{\ell} \subset \models^{g} .$$

Denn wenn für alle Punkte a in einem Modell gilt: $a \models X \Rightarrow a \models A$, dann ist es sicher auch so, daß wenn alle Punkte in einem Modell X erfüllen, dann erfüllen sie auch A. Das Beispiel

$$A \models \Box A$$

einer global aber nicht lokal gültigen Folgerung zeigt, daß die Umkehrung nicht gilt. Nur im Falle der leeren Prämissenmenge laufen beide Relationen auf dasselbe hinaus, d.h. *Formeln* – nicht Sequenzen – sind genau dann global *gültig*, wenn sie lokal gültig sind.

Die Relation $\vdash$ der Ableitbarkeit paßt also besser zu $\models$ im globalen als im lokalen Sinne. Tatsächlich koinzidiert globale Folgerung $\models^g$ in $\mathbb{K}$ mit Ableitbarkeit $\vdash$ aus Annahmen in **K**, kurz:

$$(g) \qquad \models^g_{\mathbb{K}} = \vdash_{\mathbf{K}} .$$

Auf der anderen Seite hat $\models^\ell$ die zu (Ded) passende Eigenschaft:

$$X, A \models^\ell B \Rightarrow X \models^\ell A \to B.$$

Bei diesen Beobachtungen könnten wir es bewenden lassen, denn Ableitbarkeit aus Annahmen und semantische Folgerung werden uns in diesem Kapitel nicht weiter beschäftigen – uns genügt hier der Begriff eines Theorems und der einer gültigen Formel. Es sollen aber noch kurz drei Möglichkeiten besprochen werden, sich trotz der beschriebenen Schwierigkeiten zu einem modalen Deduktionstheorem zu verhelfen. Besonders die ersten zwei Möglichkeiten sind in der Literatur häufig anzutreffen.[11]

1. Wenn uns daran liegt, dann können wir auf der Basis des axiomatischen Systems **K** eine Relation $\rhd$ einführen, welche mit lokaler Folgerung übereinstimmt:

$$(\ell) \qquad \models^\ell_{\mathbb{K}} = \rhd_{\mathbf{K}}.$$

Diese Relation $\rhd_{\mathbf{K}}$ definieren wir so:

$$(\rhd) \qquad X \rhd_{\mathbf{K}} A \text{ gdw } \exists A_1, \ldots, A_n \in X : \vdash_{\mathbf{K}} A_1 \wedge \cdots \wedge A_n \to A.$$

Daß die Relation $\rhd$ die Eigenschaft (ℓ) hat, läßt sich schnell zeigen. Denn

$$\begin{aligned} A_1, ..., A_n \models^\ell A \text{ gdw } & \forall \mathcal{M} \, \forall a \in \mathcal{M} : a \models A_1 \;\&\; \cdots \;\&\; a \models A_n \Rightarrow a \models A \\ \text{gdw } & \forall \mathcal{M} \, \forall a \in \mathcal{M} : a \models A_1 \wedge \cdots \wedge A_n \to A \\ \text{gdw } & \forall \mathcal{M} : \mathcal{M} \models A_1 \wedge \cdots \wedge A_n \to A \\ \text{gdw } & \vdash A_1 \wedge \cdots \wedge A_n \to A \text{ in } \mathbf{K} \text{ (Satz 10!)} \\ \text{gdw } & A_1, ..., A_n \rhd_{\mathbf{K}} A. \end{aligned}$$

[11] Vgl. dazu auch [81, pp. 61–64]

Die Verhältnisse faßt die folgende Tafel zusammen, wobei der Doppelpunkt : jeweils entweder $\vdash$ oder $\rhd$ vertritt.

	Ded	$A : \Box A$	$= \models^{\ell}$	$= \models^{g}$
$\vdash$	N	J	N	J
$\rhd$	J	N	J	N

2. Eine andere Möglichkeit besteht darin, an der üblichen Definition der Ableitbarkeit aus Annahmen festzuhalten, die Menge der Theoreme von **K** aber anders zu axiomatisieren. So können wir die Regel RN ersetzen durch das Schema

N. $$\Box\top\,,$$

zusammen mit der Äquivalenzregel RE. Die Menge der *Theoreme* dieses neuen axiomatischen Systems – nennen wir es $\mathbf{K}^{\top}$ – ist identisch mit **K** so wie auf p. 97 definiert. Aber die jeweiligen Relationen der Ableitbarkeit aus Annahmen sind verschieden:

$$\emptyset \vdash_{\mathbf{K}^{\top}} \,=\, \emptyset \vdash_{\mathbf{K}}, \text{ jedoch } \vdash_{\mathbf{K}^{\top}} \subset \vdash_{\mathbf{K}}\,.$$

Warum ist das so? In $\mathbf{K}^{\top}$ können wir nur eine eingeschränkte Version von RN ableiten, nämlich $\top \;/\; \Box\top$, oder äquivalent,

RN$^{\top}$. $$\frac{\vdash A}{\vdash \Box A.}$$

Im Gegensatz zu RN und MP ist diese Regel keine Ableitungsregel, sondern eine Beweisbarkeitsregel: Wenn A ein Theorem, d.h. beweisbar ist, dann ist auch $\Box A$ ein Theorem. So ist denn in $\mathbf{K}^{\top}$ die Sequenz $A \vdash \Box A$ nicht generell herleitbar und damit ist das wesentliche Hindernis für (Ded) aus dem Weg geräumt. Da das System $\mathbf{K}^{\top}$ gar keine (primitiven) modalen Regeln hat, können wir genau wie für **KL** zeigen, daß $\vdash_{\mathbf{K}^{\top}}$ sich der Eigenschaft (Ded) erfreut. [12]

3. Schließlich können wir sowohl an der üblichen Definition einer Ableitung aus Annahmen festhalten, als auch bei der Festlegung eines axiomatischen Systems auf eine Unterscheidung verschiedener Arten von Regeln verzichten. In diesem Fall können wir nach Bedingungen für X und A fragen unter denen (Ded) auch in einer Modallogik gilt, d.h. nach einem Prinzip suchen, das als eingeschränkte Variante von (Ded) erkennbar ist. [13]

[12] Eine Variante dieses Ansatzes wird in [76, Kap. 1.4] vorgestellt.

[13] Antworten auf eine der beiden Fragen können typischerweise so reformuliert werden, daß sie auch als Antwort auf die jeweils andere Frage taugen. Das gilt z.B. für das im Text vorgestellte Theorem einerseits und den Ansatz von *Fitting* in [77] andererseits.

Da in **K** nur die Regel RN der Eigenschaft (Ded) entgegensteht, ist klar, daß aus jeder modalen Variante von (Ded) folgen muß, daß (Ded) unverändert gilt, solange sich die Ableitbarkeitsbehauptung im Antezedens von (Ded) auf eine Ableitung stützen kann, die ohne RN auskommt. In allen übrigen Fällen muß die Variante die Anwendungen von RN für das Antezedens durch eine diesen Anwendungen entsprechende Abschwächung des Konsequens kompensieren. Genau dies leistet der folgende Satz.

SATZ 15. (DEDUKTIONSTHEOREM FÜR **K**, (Mints 1974) Chagrov & Zakharayaschew 1997) *Angenommen*

$$X, A \vdash B$$

aufgrund einer Ableitung, in welcher RN n-mal ($n \geq 0$) auf Formeln angewandt wurde, die in der Ableitung von A abhängig sind. Dann

$$X \vdash \Box^0 A \wedge \cdots \wedge \Box^n A \to B.$$

(Hier steht $\Box^n$ für die n-fache Iteration des $\Box$-Operators; also $\Box^0 A = A, \Box^1 A = \Box A$, ... , $\Box^n A = \Box\Box^{n-1} A$.)

BEWEIS. Induktion über die Länge k einer Ableitung von B aus (X, A). Im Fall $k = 1$ gibt es keine Anwendungen von RN und so argumentieren wir wie für das nicht-modale Deduktionstheorem. ($X \vdash^k A$ deute eine Ableitung von A aus X in k Schritten an.)

Im Induktionsschritt nehmen wir an (IA), daß der Satz für alle Ableitungen kürzer als k gilt. Wir unterscheiden nun zwei Fälle.

Erster Fall: B wurde durch einen MP-Schritt gewonnen. Dann gibt es $i, j < k$ mit

$$X, A \vdash^i C \text{ und } X, A \vdash^j C \to B,$$

wobei RN ℓ- bzw. m-mal ($\ell, m < n$) auf Formeln abhängig von A angewandt wurde. Wir wenden die IA an und erhalten

$$X \vdash \Box^0 A \wedge \cdots \wedge \Box^\ell A \to C \text{ und } X \vdash \Box^0 A \wedge \cdots \wedge \Box^m A \to (C \to B),$$

woraus die gewünschte Sequenz folgt.

Zweiter Fall: B wurde durch einen RN-Schritt gewonnen. Dann gibt es $i < k$ mit

$$X, A \vdash^i C \text{ und } B = \Box C$$

Wir wenden die IA an und erhalten

$$X \vdash \Box^0 A \wedge \cdots \wedge \Box^m A \to C$$

unter $m < n$ Anwendungen von RN auf von A abhängigen Formeln. Eine Anwendung von RK ergibt

$$X \vdash \Box\Box^0 A \wedge \cdots \wedge \Box\Box^m A \to \Box C, \text{ d.h.}$$
$$X \vdash \Box^1 A \wedge \cdots \wedge \Box^{m+1} A \to B,$$

woraus die gewünschte Sequenz folgt. ∎

Korrollar 16. *In* **K** *gilt: Wenn $X, A \vdash B$ aufgrund einer Ableitung, in der RN nicht angewandt wurde ($n = 0$ im Satz), dann $X \vdash A \to B$.*

Um den Satz auf eine Sequenz $X, A \vdash B$ anwenden zu können, reicht es nicht – wie im Falle des klassischen Deduktionstheorems – zu wissen, daß es eine Ableitung von B aus (X, A) gibt. Vielmehr müssen wir eine Ableitung vorlegen und prüfen, wie oft RN auf von A abhängige Formeln in der Ableitung angewandt wurde. D.h. wir brauchen mehr und vor allem eine andere Art von Information als für das klassische Deduktionstheorem. Das ändert sich, wenn wir Erweiterungen von **K** betrachten. Es sei **KT4** – auch bekannt als **S4** – die Erweiterung von **K** um die Schemata T und 4 (siehe p. 87).

Satz 17. (Deduktionstheorem für **KT4**) *Wenn $X, A \vdash B$, dann $X \vdash \Box A \to B$.*

Beweis. Angenommen $X, A \vdash B$. Dann gibt es eine Ableitung von B aus (X, A), welche die Bedingung im Antezedens des Satzes 15 erfüllt. Somit

$$X \vdash \Box^0 A \wedge \cdots \wedge \Box^n A \to B. \tag{1}$$

Aus 4 und RM erhalten wir die Implikationskette

$$\vdash \Box^1 A \to \cdots \to \Box^n A \tag{2}$$

und somit aus (1)

$$X \vdash A \wedge \Box A \to B. \tag{3}$$

Mit T, $\vdash \Box A \to A$, folgt dann aus (3) die gewünschte Sequenz $X \vdash \Box A \to B$. ∎

7. Erweiterungen von K

Wir wollen mit $\mathbf{K}^+$ variabel normale Modallogiken bezeichnen, also Erweiterungen von $\mathbf{K}$ unter der Bedingung, daß diese unter MP, RN und gleichförmiger Einsetzung abgeschlossen sind. Wenn X ein Formel- oder Regelschema ist, dann sei $\mathbf{K} + X$ die normale Erweiterung von $\mathbf{K}$ um das Schema X; wir schreiben dafür manchmal auch $\mathbf{KX}$.

Man beachte, daß $\mathbf{K} + X$ nicht dieselbe Menge ist wie $\mathbf{K} \cup \{A : A \text{ ist eine Instanz von } X\}$. Diese letztere Menge ist weder unter MP noch unter RN abgeschlossen, d.h. keine normale Modallogik. Die Operation $+$ der normalen Erweiterung einer Modallogik ist also stärker als die einfache Mengenoperation $\cup$.

Auf der einen Seite haben wir also normale Modallogiken, auf der anderen Seite Kripke-Rahmen – im folgenden einfach "Rahmen" genannt – und deren Theorien. Zur Erinnerung:

(Def. *Th*) $$Th(\mathbb{R}) = \{A \in \mathcal{L}_\Box : \mathbb{R} \models A\}.$$

Da die kleinste normale Logik $\mathbf{K}$ die Theorie der Kripke-Rahmen im *allgemeinsten* Sinne abgibt, so liegt die Vermutung nahe, daß jede normale Logik die Theorie einer Klasse von Kripke-Rahmen in einem *eingeschränkteren* Sinne ist. Wenn das so wäre, dann könnten wir auch sagen: Eine Modallogik ist genau dann normal, wenn sie die Theorie einer Klasse von Rahmen ist. Das ist nicht so, wie wir gleich sehen werden. Festhalten können wir jedoch die folgende

Beobachtung 18. *Für jede Rahmenklasse $\mathbb{R} \subseteq \mathbb{K}$ ist $Th(\mathbb{R})$ eine normale Modallogik.*

Beweis. Allgemein gilt, daß wenn $\mathbb{R} \subseteq \mathbb{K}$, dann $Th(\mathbb{K}) \subseteq Th(\mathbb{R})$. Da $Th(\mathbb{K}) = \mathbf{K}$, ist $Th(\mathbb{R})$ also eine Erweiterung von $\mathbf{K}$. Ferner ist Gültigkeit in Rahmen unter MP, RN und gleichförmiger Einsetzung abgeschlossen. Also ist $Th(\mathbb{R})$ eine normale Erweiterung von $\mathbf{K}$. ∎

Die Basislogik $\mathbf{K}$ ist offenbar zu schwach, um die logischen Wahrheiten, die wir unter den oben genannten Beispielinterpretationen von $\Box$ erwarten dürfen, einzufangen. Dieses System kann nur als Ausgangspunkt dienen für den Aufbau von Systemen, die eher als $\mathbf{K}$ dafür geeignet sind, eine Logik der Notwendigkeit abzugeben. Eines der ersten Postulate, die in diesem Sinne gute Kandidaten sind, wäre

T. $$\Box A \to A.$$

Aber auch z.B.

4. $$\Box A \to \Box\Box A$$

sieht nicht schlecht aus. Weitere prominente Schemata, die wir in diesem Kapitel in den Blick nehmen werden, sind diese:

D. $$\Box A \to \Diamond A;$$

B. $$A \to \Box\Diamond A;$$

5. $$\Diamond A \to \Box\Diamond A.$$

Die Erweiterung von **K** um, beispielsweise, die Schemata T und 4, $\mathbf{K} + (\mathrm{T}, 4)$, bezeichnen wir auch kurz mit **KT4**. Diese Logik taucht schon als **S4** in C.I. Lewis' Reihe von Modallogiken auf. Uns interessiert natürlich die Frage, welche Art von Notwendigkeit **KT4** beschreibt. Dazu wäre es hilfreich zu wissen, welche strukturierten Möglichkeitsräume, d.h. welche Rahmenklasse **KT4** richtig und vollständig beschreibt. Mit anderen Worten, wir fragen nach einer Rahmenklasse $\mathbb{R}$ so, daß $Th(\mathbb{R}) = \mathbf{KT4}$. Diese Rahmenklasse werden wir bald beschreiben. Wir wollen die Frage aber etwas allgemeiner angehen, indem wir zunächst über das grundsätzliche Verhältnis zwischen Rahmenbedingungen und modallogischen Formeln nachdenken.

8. Determination, Definition und Korrespondenz

Im Kapitel II über Zeitlogik haben wir gesehen, wie sich zeitlogische Formeln in die Prädikatenlogik übersetzen lassen. Dasselbe gilt natürlich für die Formeln der strukturell gleichen Sprache $\mathcal{L}_\Box$. So ist die Aussage, daß die Formel $\Box P \to P$ ($P \in \mathrm{ATM}$) am Punkt a eines Modells wahr ist, genau dann wahr, wenn
$$\forall y(Ray \Rightarrow y \models P) \Rightarrow a \models P.$$

Die Eigenschaft eines Punktes y, P in einem gegebenen Modell wahr zu machen, können wir auch mit Py notieren, und $\Rightarrow$ verstehen wir im Sinne der materialen Implikation $\to$. So erhalten wir

t(a) $$\forall y(Ray \to Py) \to Pa.$$

Ebenso können wir die Wahrheit von $\Box P \to P$ an allen Punkten eines Modells, also Wahrheit im Modell, in die Sprache der *ersten* Stufe übersetzen:

t($\mathcal{M}$) $$\forall x(\forall y(Rxy \to Py) \to Px).$$

Wie steht es nun mit Wahrheit in allen Modellen, d.h. Gültigkeit in einem Rahmen $\mathcal{K}$? Jetzt quantifizieren wir über wechselnde Interpretationen von P, d.h. wechselnde Eigenschaften von Punkten. So erhalten wir als Übersetzung der Aussage, daß $\Box P \to P$ in *allen* Modellen wahr sei, eine Formel *zweiter* Stufe:

$$\mathrm{t}(\mathcal{K}) \qquad \forall\phi\forall x(\forall y(Rxy \to \phi y) \to \phi x).$$

Das ist nicht weiter überraschend: Jede modallogische Formel korrespondiert per Übersetzung mit einer prädikatenlogischen Formel. Überraschend ist vielmehr, daß sich manchmal die Quantifikation zweiter Stufe über Interpretationen ($\forall\phi$...) eliminieren läßt und so eine Formel erster Stufe entsteht, welche nur auf die Rahmenkomponenten von Modellen Bezug nimmt: Die Modellbeschreibung zweiter Stufe erweist sich als äquivalent zu einer Rahmenbeschreibung erster Stufe. Wenn wir z.B. in $\mathrm{t}(\mathcal{K})$ für ϕ die Eigenschaft $Rx(\ )$ einsetzen, dann erhalten wir

$$\forall x(\forall y(Rxy \to Rxy) \to Rxx),$$

woraus offensichtlich $\forall x(Rxx)$ folgt. In umgekehrter Richtung nehmen wir an, daß (1) $\forall x(Rxx)$ und (2) $\forall y(Ray \to Py)$ (mit Punkt a und Eigenschaft P beliebig gewählt). Das zu zeigende Pa folgt dann unmittelbar, indem wir a in (1) für x und in (2) für y einsetzen. Zusammengefaßt: Die Standardübersetzung der Aussage, daß $\Box P \to P$ in allen Modellen auf (W, R) wahr ist, $\mathrm{t}(\mathcal{K})$, ist äquivalent zu $\forall x(Rxx)$.

Bemerkung. An dieser Stelle entsteht natürlich die Frage nach einer allgemeinen Charakterisierung solcher (Modelle beschreibenden) Formeln zweiter Stufe, die zu einer (Rahmen beschreibenden) Formel erster Stufe äquivalent sind. Eine Antwort auf diese Frage gibt der *Satz von Sahlqvist*, den wir hier aber nicht behandeln werden.[14] Es seien hier nur die wichtigsten Ergebnisse zusammengefaßt. Modale Axiome lassen sich generell in die Sprache zweiter Stufe übersetzen. Wenn die Übersetzung eine sogenannte Sahlqvist-Formel ist – und diese Bedingung ist entscheidbar –, dann läßt sich effektiv eine dem Axiom entsprechende Formel erster Stufe finden (Sahlqvist-van Benthem-Algorithmus). Wenn die Übersetzung keine Sahlqvist-Formel ist, dann ist die Frage zunächst offen, ob es die gewünschte Formel erster Stufe gibt; tatsächlich ist diese Frage nicht entscheidbar (Satz von Chagrova [51]). Gibt es nun Bedingungen erster Stufe, die nicht äquivalent zu einer Sahlqvist-Formel sind aber dennoch

[14] Siehe dazu [31, Kap. 3–4].

einem modalen Axiom entsprechen? Ein solches Beispielpaar wird in [31, pp. 167–169] vorgeführt: Die Übersetzung des McKinsey-Axioms

M. $$\Box\Diamond A \to \Diamond\Box A$$

ist ebensowenig eine Sahlqvist-Formel wie die Übersetzung der Konjunktion M $\wedge$ $(\Box A \to \Box\Box A)$.[15] Die Konjunktion korrespondiert jedoch mit dieser Bedingung erster Stufe:

$$\forall x \exists y (Rxy \wedge \forall z (Ryz \to z = y)).$$

(Ende der Bemerkung.)

Klassen von Kripke-Rahmen können im Prinzip auf zweierlei Weise festgelegt werden:

1. Wir können Rahmen unter eine direkte Bedingung stellen. Im folgenden verstehen wir darunter immer eine Bedingung, die sich in der Sprache der ersten Stufe formulieren läßt. Für eine solche Bedingung (Bed) notieren wir mit $\mathit{Bed}(\mathbb{K})$ die Klasse derjenigen Rahmen in $\mathbb{K}$, welche (Bed) erfüllen.
2. Wir können Rahmen unter eine indirekte Bedingung stellen, nämlich die Bedingung, daß in ihnen die Theoreme einer bestimmten normalen Modallogik **L** gültig seien. Für eine Logik **L** notieren wir mit $\mathit{Rm}(\mathbf{L})$ die Klasse der diese Bedingung erfüllenden Rahmen, der *Rahmen für* **L**, wie wir auch sagen werden.

$$\mathit{Rm}(\mathbf{L}) = \{\mathcal{K} \in \mathbb{K} : \mathcal{K} \models \mathbf{L}\}.$$

Aus der Definition geht unmittelbar hervor, daß

$$\mathit{Th}(\mathit{Rm}(\mathbf{L})) = \mathbf{L}.$$

Man beachte ferner, daß wenn S ein Formelschema ist, dann

$$\mathit{Rm}(S) = \mathit{Rm}(\mathbf{K} + S)$$

(und $\mathit{Th}(\mathit{Rm}(S)) = \mathbf{K} + S$). Denn jeder Rahmen ist schon ein Rahmen für das Basissystem **K**.

Wenn nun $\mathit{Th}(\mathit{Bed}(\mathbb{K})) = \mathbf{L}$, dann legt die Rahmenbedingung (Bed) gewissermaßen die Logik **L** fest; wenn andererseits $\mathit{Rm}(\mathbf{L}) = \mathit{Bed}(\mathbb{K})$, dann legt die Logik **L** die Bedingung (Bed) fest. Im ersten Fall sagt man auch, daß die Bedingung (bzw. die Rahmenklasse, die unter der Bedingung steht) die Logik *determiniert*; im zweiten Fall sagt man, daß die Logik die Bedingung *definiert*. Die folgende Definition faßt das zusammen:

---

[15] Van Benthem (in [30, p. 12]) beobachtet, daß M äquivalent ist zu $\neg\Box(\Diamond A \wedge \Diamond\neg A)$ und bietet eine interessante Glossierung an: “Nichts ist absolut relativ”.

DEFINITION 19. Es sei $\mathbf{L}$ ($\supseteq \mathbf{K}$) eine normale Modallogik, $\mathbb{L}$ ($\subseteq \mathbb{K}$) sei eine Rahmenklasse, (Bed) sei eine Bedingung erster Stufe auf Rahmen.

1. Die Rahmenklasse $\mathbb{L}$ *determiniert die Logik* $\mathbf{L}$ gdw $Th(\mathbb{L}) = \mathbf{L}$.
2. Die Logik $\mathbf{L}$ *definiert die Rahmenklasse* $\mathbb{L}$ gdw $Rm(\mathbf{L}) = \mathbb{L}$.
3. Die Bedingung (Bed) *determiniert das Schema* S gdw $Th(Bed(\mathbb{K})) = \mathbf{K} + S$.
4. Das Schema S *definiert die Rahmenbedingung* (Bed) gdw $Rm(S) = Bed(\mathbb{K})$.

Wenn ein Schema eine Rahmenbedingung definiert, dann wollen wir auch sagen, daß Schema und Bedingung miteinander *korrespondieren.* Daß eine Bedingung (Bed) ein Schema S determiniert, haben wir zuvor auch so ausgedrückt: Die Logik $\mathbf{K} + S$ ist richtig und vollständig im Hinblick auf alle Rahmen, welche die Bedingung (Bed) erfüllen.

Nach der Definition stehen ein Schema S und eine Rahmenbedingung (Bed) genau dann in der Korrespondenzrelation, wenn die Rahmen für S genau die (Bed)-Rahmen sind, d.h.

(Korr) $$\forall \mathcal{M} \in \{(\mathcal{K}, I) : \mathcal{K} \in \mathbb{K}\} : \mathcal{M} \models S \text{ gdw } Bed(\mathcal{K}).$$

Für den Nachweis einer solchen Korrespondenz genügt es also einen beliebigen Punkt a in einem beliebigen Modell $(\mathcal{K}, I)$ zu betrachten und zu zeigen, daß

$$a \models S \text{ gdw } \mathcal{K} \text{ die Bedingung (Bed) erfüllt.}$$

Wir wollen das an einem einfachen Beispiel veranschaulichen.

BEOBACHTUNG 20. *Die Bedingung (Refl), Raa, korrespondiert mit dem Schema T,* $\Box A \to A$.

BEWEIS. *Rechts nach links.* Wir betrachten einen Punkt a in einem reflexiven Rahmen und zeigen, daß $a \models \Box A \to A$:

(1)	$a \models \Box A$	Annahme (zz $a \models A$)
(2)	$\forall x : Rax \Rightarrow x \models A$	1, ($\Box$)
(3)	Raa	Reflexivität!
(4)	$a \models A$	2,3

Links nach rechts. Wir kontraponieren (Korr) und betrachten einen beliebigen nichtreflexiven Rahmen $\mathcal{K}$. (Wir müssen nun ein Modell auf $\mathcal{K}$ finden, welches an einem Punkt eine Instanz von T falsifiziert.) Sei a ein nichtreflexiver Punkt in $\mathcal{K}$. Das Modell $(\mathcal{K}, I)$ können wir so wählen, daß für ein Atom P, $I(P, a) = 0$ und für alle x mit Rax, $I(P, x) = 1$. Dann haben wir $a \models \Box P$ während $a \not\models P$, wie gewünscht. (Alternativ würde jedes Modell mit $I(P, a) = 0$ und $I(P, x) = 1$ mit x aus einer beliebigen Obermenge von $R(a) = \{x : Rax\}$ demselben Zweck dienen.) ∎

Wir wissen nun, daß T und (Refl) miteinander korrespondieren; oder anders ausgedrückt: T definiert (Refl), bzw. die Logik **KT** definiert *Refl*($\mathbb{K}$). Dürfen wir daraus schließen, daß die Klasse der reflexiven Rahmen die Logik **KT** *determiniert*? Wir fragen hier nach der Implikation

$$(*) \qquad Rm(\mathbf{KT}) = \mathit{Refl}(\mathbb{K}) \;\Rightarrow\; Th(\mathit{Refl}(\mathbb{K})) = \mathbf{KT}.$$

Nun ist es so, daß das Konsequens von $(*)$ in diesem Fall wahr ist: die Klasse der reflexiven Rahmen determiniert **KT**. Tatsächlich jedoch können wir *im allgemeinen* aus der Korrespondenz, d.h. der Definition einer Rahmenklasse durch eine Logik, nicht auf die Determination der Logik durch diese Rahmenklasse schließen. Schon das Beispiel von **KT** und (Refl) verdeutlicht die Problematik eines solchen Schlusses.

Um zu beweisen, daß **KT** durch die Klasse der reflexiven Rahmen determiniert ist, führen wir ein Vollständigkeitsargument mit Hilfe des kanonischen Modells $\mathcal{M}^{\mathbf{KT}}$ für **KT**. Nachdem wir uns davon überzeugt haben, daß $\mathcal{M}^{\mathbf{KT}} \models \Box A \to A$, verbleibt zu zeigen, daß $\mathcal{M}^{\mathbf{KT}}$ reflexiv ist. Letzteres folgt aber nicht aus der bloßen Korrespondenz $Rm(\mathbf{KT}) = \mathit{Refl}(\mathbb{K})$. Aus der Korrespondenz können wir nur schließen, daß

$$\text{Wenn } \forall I : (\mathcal{K}^{\mathbf{KT}}, I) \models \Box A \to A, \text{ dann } \mathit{Refl}(\mathcal{K}^{\mathbf{KT}}).$$

Wir haben die kanonische Interpretation $I^{\mathbf{KT}}$ so definiert, daß sie das T-Schema in $\mathcal{K}^{\mathbf{KT}}$ verifiziert; aber wie steht es mit anderen Interpretationen auf $\mathcal{K}^{\mathbf{KT}}$? Solange wir solche Interpretationen, die das Schema falsifizieren, nicht ausschließen können, bleibt trotz Korrespondenz die Frage offen, ob $\mathcal{M}^{\mathbf{KT}}$ ein nichtreflexives Modell für **KT** ist. Genau diese Möglichkeit gilt es auszuschließen, wenn wir von der Korrespondenz zur Determinierung übergehen möchten.

Es ist übrigens nicht schwierig Modelle zu finden, in denen T gilt, R jedoch nicht reflexiv ist. Man betrachte zum Beispiel $(\{a, b\}, R, I)$ mit Rab und Rba. Für alle Atome P sei $I(P, a) = 0 = I(P, b)$. Eine Induktion über A zeigt, daß $\Box A \to A$ an jedem Punkt dieses nichtreflexiven Modells wahr ist. Das liegt natürlich daran, daß wir I so definiert haben, daß T erzwungen wird, ohne daß R im Modell reflexiv sein muß. Im kanonischen Modell für eine Logik definieren wir I auch so, daß wir die Wahrheit der Theoreme im Modell erzwingen. Nichts schließt bisher aus, daß wir das tun können ohne auch die korrespondierenden Eigenschaften der Relation R zu erzwingen.

Die Korrespondenz zwischen (Refl) und T verhilft uns also nur dann zum Determinationsresultat, wenn wir zeigen können, daß T im kanonischen

Rahmen $\mathcal{K}^{\mathbf{KT}}$ gültig ist (und nicht bloß im kanonischen Modell $\mathcal{M}^{\mathbf{KT}}$ wahr ist). Wenn das so ist, dann – so wollen wir sagen – ist T *kanonisch.*

- Eine *Formel* A ist kanonisch, wenn für jede normale Modallogik $\mathbf{L}$ mit $A \in \mathbf{L}$ gilt: A ist gültig im kanonischen Rahmen für $\mathbf{L}$.
- Eine *Logik* $\mathbf{L}$ ist kanonisch, wenn jedes Theorem von $\mathbf{L}$ kanonisch ist.
- Ein *Schema* S ist *für eine Rahmenbedingung* (Bed) kanonisch, wenn (a) S in jedem (Bed)-Rahmen gültig ist, und (b) für jede normale modale Logik $\mathbf{L}$ mit $A \in \mathbf{L}$ gilt: der kanonische Rahmen für $\mathbf{L}$ erfüllt (Bed).

Um unter der Vorraussetzung der Korrespondenz zu zeigen, daß ein Schema kanonisch für eine Bedingung ist, genügt es offenbar, sich von (b) zu überzeugen, d.h. die Bedingung für den $\mathbf{L}$-kanonischen Rahmen einer beliebigen normalen Modallogik $\mathbf{L}$, die das Schema enthält, nachzuweisen. Der Beweis der nächsten Beobachtung führt das beispielhaft vor.

BEOBACHTUNG 21. *T ist kanonisch für (Refl).*

BEWEIS. Wir haben unter 20 schon gezeigt, daß T und (Refl) korrespondieren. Wie betrachten nun das kanonische Modell $\mathcal{M}^{\mathbf{L}}$ für eine beliebige Logik $\mathbf{L}$ mit T als Theorem und nehmen an, das Modell enthalte einen irreflexiven Punkt a, also *nicht* Raa. (Die Punkte im kanonischen Modell, obgleich Formelmengen, bezeichnen wir fortan mit Kleinbuchstaben.) Nach der Definition von R im kanonischen Modell, ist dann $a^{\Box}$ nicht in a enthalten, d.h. es gibt eine Formel A in $a^{\Box}$ und nicht in a. $A \in a^{\Box}$ bedeutet: $\Box A \in a$. Aber $\mathcal{M}^{\mathbf{L}}$ ist charakteristisch für $\mathbf{L}$ (M6) und also haben wir $\Box A \to A \in a$. Ferner ist a unter MP abgeschlossen. Also ist $A \in a$, im Widerspruch zu unserer Annahme. Die Rahmenkomponente $\mathcal{K}^{\mathbf{L}}$ des kanonischen Modells ist also reflexiv. Aus der Korrespondenz (Kor)(rechts nach links) von T mit (Refl), dürfen wir daher schließen, daß $\mathcal{K}^{\mathbf{L}} \models$ T, der kanonische Rahmen für $\mathbf{L}$ also ein Rahmen für T ist. ■

Da T kanonisch für (Refl) ist, wissen wir nun insbesondere, daß das kanonische Modell für $\mathbf{KT}$, welches beliebige Nicht-Theoreme von $\mathbf{KT}$ falsifiziert, wie gewünscht reflexiv ist. So folgt unmittelbar aus den Beobachtungen 20 und 21 das

KOROLLAR 22. *Die Logik* $\mathbf{KT}$ *ist richtig und vollständig bezüglich der (determiniert durch die) Klasse reflexiver Rahmen.*

Das gerade erprobte Rezept für Determinationsresultate mit der kanonischen Modell-Methode läßt sich offensichtlich nicht nur verallgemeinern, sondern auch unter eine schwächere Bedingung stellen. Tatsächlich brauchen wir für den Richtigkeitsteil der Determination (a) nur eine Hälfte der Korrespondenz; und für den Vollständigkeitsteil genügt es, sich davon zu überzeugen, daß (b) das kanonische Modell die fragliche Bedingung erfüllt. Diese Beobachtung führt uns zur Einführung einer weiteren Relation:

- Das Schema S *korrespondiert kanonisch* mit der Bedingung (Bed) gdw
  (a) $S \in Th(Bed(\mathbb{K}))$, und
  (b) für jede normale modale Logik $\mathbf{L}$ mit $S \in \mathbf{L}$ hat das kanonische Modell $\mathcal{M}^{\mathbf{L}}$ die Eigenschaft (Bed).

Es ist diese Relation, die im Beweis der Richtig- und Vollständigkeit einer Modallogik im Hinblick auf eine Rahmenklasse zum Zuge kommt; praktischerweise setzen wir dabei für die Variable $\mathbf{L}$ das System $\mathbf{K} + S$ ein. Man beachte, daß viele Autoren mit "Korrespondenz" die Relation der *kanonischen* Korrespondenz meinen![16]

Satz 23. *Es sei ein S ein Formelschema in $\mathcal{L}_\Box$ und (Bed) eine Rahmenbedingung. Wenn (a) $S \in Th(Bed(\mathbb{K}))$ und (b) das kanonische Modell für $\mathbf{K} + S$ die Bedingung (Bed) erfüllt, dann determiniert die Klasse der Rahmen unter der Bedingung (Bed) die Logik $\mathbf{K}+S$: $\mathbf{K}+S = Th(Bed(\mathbb{K}))$.*

Der Satz schließt natürlich auch den Fall ein, daß $\mathbf{K}$ gleich um mehrere Axiome erweitert wird; S möge dann für die Konjunktion dieser Axiome und Bed für die Konjunktion der den einzelnen Axiomen entsprechenden Bedingungen stehen. Etwas allgemeiner halten wir das in einem Korollar fest.

Korollar 24. *Es seien $S_1, ..., S_n$ Formelschemata und $Bed_1, ..., Bed_n$ Rahmenbedingungen so, daß jedes Paar (S_i, Bed_i) $(1 \leq n \leq n)$ jeweils die Bedingungen (a) und (b) des Satzes erfüllt. Dann gilt*

$$\mathbf{K} + S_1 + \cdots S_n = Th(Bed_1(\mathbb{K}) \cap \cdots \cap Bed_n(\mathbb{K})).$$

Wir betrachten nun einige wichtige Korrespondenzen. Tafel 1a enthält eine Liste von Rahmenbedingungen. In der rechten Spalte sind die Namen der korrespondierenden Schemata angegeben, welche in Tafel 1b zusammengestellt sind. All diese Schemata sind kanonisch, weshalb die Tafeln kanonische Korrespondenzen angeben, aus denen gleich Determinationsresultate ablesbar sind. Beides müssen wir noch zeigen: daß es sich um Korrespondenzen handelt und daß die Schemata kanonisch sind.[17]

[16] Im Zweifel schaue man auf die Beweise von "Korrespondenz"-Beobachtungen.
[17] Weiter unten (p. 150) werden wir ein Schema kennenlernen, das Löb-Schema W, welches nicht kanonisch ist.

Bedingungen

(Refl)exiv	Raa	T
(Trans)itiv	$Rab \,\&\, Rbc \Rightarrow Rac$	4
(Sym)metrisch	$Rab \Rightarrow Rba$	B
(Eukl)idisch	$Rab \,\&\, Rac \Rightarrow Rbc$	5
(Erw)eiterbar	$\forall a \exists b : Rab$	D
(Dicht)	$Rab \Rightarrow \exists c : Rac \,\&\, Rcb$	4_c
(Konv)ergent	$Rab \,\&\, Rac \Rightarrow \exists d : Rbd \,\&\, Rcd$	G
(Unv)erbunden	$R = \emptyset$	Ver

Schemata

(Refl)exiv	$\Box A \to A$	$A \to \Diamond A$	T
(Trans)itiv	$\Box A \to \Box\Box A$	$\Diamond\Diamond A \to \Diamond A$	4
(Sym)metrisch	$A \to \Box\Diamond A$	$\Diamond\Box A \to A$	B
(Eukl)idisch	$\Diamond A \to \Box\Diamond A$	$\Diamond\Box A \to \Box A$	5
(Erw)eiterbar	$\Box A \to \Diamond A$		D
(Dicht)	$\Box\Box A \to \Box A$	$\Diamond A \to \Diamond\Diamond A$	4_c
(Konv)ergent	$\Diamond\Box A \to \Box\Diamond A$		G
(Unv)erbunden	$A \to \Box A$	$\Diamond A \to A$	T_c

Ob wir die eine oder andere Eigenschaft von R für richtig halten, hängt von der Art der Notwendigkeit bzw. Möglichkeit ab, die wir im Auge haben. In jedem Fall schränken wir aber durch solche Eigenschaften die Klasse der betrachteten Modelle ein und erweitern damit die Chancen von Schemata auf logische Gültigkeit.

SATZ 25. (KORRESPONDENZ) *Für die Paare (Bed,S) in der obigen Tafel gilt: Ein Rahmen $\mathcal{K}$ erfüllt die Bedingung (Bed) gdw $\mathcal{K} \models S$ (d.h. $Rm(\mathbf{K} + S) = Bed(\mathbb{K})$).*

BEWEIS. Für die LR-Bedingung ist in einem beliebigen Modell auf Rahmen, welche die Bedingung erfüllen, die Wahrheit des Axioms an einem beliebig gewählten Punkt nachzuweisen. Die Ausführung sei dem Leser überlassen.

Die RL-Bedingung beweisen wir, indem wir ein Modell so konstruieren auf einem beliebigen Rahmen, der die Bedingung nicht erfüllt, daß eine Instanz des Axioms an einem Punkte falsch wird. Notation: $R(a) = \{x : Rax\}$ und $R(a) \models A$ gdw $x \models A$ für alle $x \in R(a)$. Die Korrespondenz von T mit (Refl) wurde oben schon unter 20 bewiesen.

Ad (Trans, 4): Wir betrachten einen Rahmen mit Punkten a, b, c so, daß Rab und Rbc, jedoch nicht Rac. Für I fordern wir $I(P, c) = 0$, sowie $R(a) \models P$. Dann $a \models \Box P$. Da Rbc, ist $b \not\models \Box P$, weshalb $a \not\models \Box\Box P$.

Ad (Sym, B): In einem Rahmen gebe es Punkte a und b mit Rab, jedoch nicht Rba. Wir setzen $I(P, a) = 1$ sowie $R(b) \models \neg P$. Nun besagt $a \models \Box\Diamond P$, daß $\forall x : Rax \Rightarrow \exists y : Rxy \;\&\; y \models P$. Danach müßte es insbesondere für b einen Punkt y geben mit $y \models P$, was nach Voraussetzung ausgeschlossen ist.

Ad (Eukl, 5): Ein Rahmen enthalte Punkte a, b, c mit Rab und Rac, jedoch weder Rbc noch Rcb. Wir setzen $I(P, c) = 1$ sowie $R(b) \models \neg P$. Dann $a \models \Diamond P$ aufgrund von Rac. Nun behauptet $a \models \Box\Diamond P$, daß es von b und von c ausgehend erreichbare Punkte gibt, die P wahr machen. Aber für b ist das nach der Voraussetzung nicht der Fall.

Ad (Erw, D): Hier betrachten wir einen Rahmen mit einem R-Endpunkt a. Dann gilt $a \models \Box P$ trivialerweise; und da es keinen Punkt gibt, der von a aus erreichbar ist, so gilt nicht $a \models \Diamond P$. (*NB*: Wir haben hier keine Annahmen über die Bewertung von Formeln gemacht. In Rahmen mit Endpunkten werden *alle* Instanzen von D falsch.)

Ad (Dicht, 4c): In einem Rahmen gebe es Punkte a und b mit Rab und ohne einen Punkt c zwischen a und b (d.h. nicht Rac und Rcb). Ferner setzen wir (a) $I(P, b) = 0$ und (b) für alle x mit Rax: $R(x) \models P$. Dann haben wir aufgrund von (b) $a \models \Box\Box P$. Nun behauptet $a \models \Box P$, daß insbesondere $b \models P$, was nach (a) falsch ist. Man beachte, daß im (ausgeschlossenen) Falle eines Punktes x zwischen a and b, die Festsetzungen für b und $R(x)$ inkonsistent wären. Wir hätten dann $b \models P$ und $b \models \neg P$.

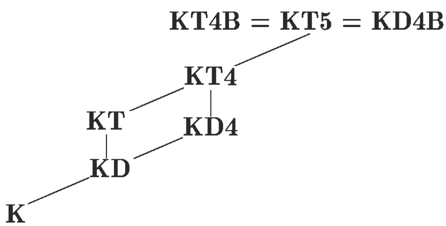
Ad (Konv, G): Wir betrachten einen Rahmen mit Rab, Rac, Rcd, jedoch nicht Rbd. Wir setzen $R(b) \models P$ und $R(c) \models \neg P$. Dann gilt $b \models \Box P$ und also (Rab!) $a \models \Diamond\Box P$. Jedoch ist P an keinem Punkt in $R(c)$ wahr; daher $c \not\models \Diamond P$ und also (Rac!) $a \not\models \Box\Diamond P$. Man beachte, daß im (ausgeschlossenen) Falle Rbd die Festsetzungen für $R(b)$ und $R(c)$ inkonsistent wären: Wir hätten dann $d \models P$ und $d \models \neg P$.

Ad (Unv, T_c): Übung. ∎

Die Tafel auf p. 121 gibt gewissermaßen einen Baukasten ab, aus dem wir uns bedienen können, um Logiken und dazu passende Rahmen zu basteln. Auf der Basis eines umfangreicheren Baukastens listet Chellas [53, pp. 285f.] eine Auswahl (!) von 142 Erweiterungen von **K** auf. Von Bedeutung sind aber nur eine Handvoll solcher Systeme. Einige Logiken haben traditionelle Namen, die wir rechts anführen:

$$\begin{aligned} \mathbf{KD} &= \mathbf{SDL} \\ \mathbf{KT} &= \mathbf{T} \\ \mathbf{KT4} &= \mathbf{S4} \\ \mathbf{KT4G} &= \mathbf{S4.2} \\ \mathbf{KT5} &= \mathbf{S5} \end{aligned}$$

KD (= **SDL**) ist das sogenannte Standardsystem Deontischer Logik (zuerst in von Wright [287]; siehe auch die Einleitung in [136]). **KT4** (= **S4**) ist ein System, welches oft favorisiert wird sowohl als die Logik epistemischer als auch metaphysischer Notwendigkeit. **KT4B** (= **S5**), welches äquivalent ist zu **KT5**, gilt, wie eingangs erklärt, als bester Kandidat für die Logik logischer Notwendigkeit. Auf das sogenannte intermediäre Systeme **S4.2** werden wir später zurückkommen. Einen Eindruck von der relativen Stärke der Systeme gibt die folgende Tafel. (Schwächere Systeme unter stärkeren; alle Kanten geben echte Inklusionsverhältnisse an.)



Manche Kombinationen von **K** mit verschiedenen Axiomen haben denselben Effekt. So läßt sich zum Beispiel das System **S5** auf verschiedene Weisen axiomatisieren.

Beobachtung 26. **KT5** = **KT4B** = **KD4B** = **KDB5** (= **S5**).

Beweis. Zu beweisen ist eine zyklische Inklusionskette, zum Beispiel

$$\mathbf{KT5} \subseteq \mathbf{KT4B} \subseteq \mathbf{KDB4} \subseteq \mathbf{KDB5} \subseteq \mathbf{KT5}.$$

Für diese Kette genügt es zu zeigen, daß (a) $\mathbf{KT4B} \vdash 5$, (b) $\mathbf{KD4B} \vdash \mathrm{T}$, (c) $\mathbf{KDB5} \vdash 4$, sowie (d) $\mathbf{KT5} \vdash \mathrm{D}$ und $\mathbf{KT5} \vdash \mathrm{B}$. (Übung.) ∎

Aus der Äquivalenz der Kombination von Schemata zum System **S5** folgt, aufgrund des Korrespondenzsatzes, auch die Äquivalenz der Kombination der korrespondierenden Bedingungen. Also:

$$\left\{\begin{matrix}\text{Refl}\\\text{Trans}\\\text{Sym}\end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix}\text{Refl}\\\text{Eukl}\end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix}\text{Erw}\\\text{Trans}\\\text{Sym}\end{matrix}\right\}.$$

(Umgekehrt stellt der direkte Nachweis der Äquivalenz der drei Bedingungskombinationen einen alternativen Beweis der Beobachtung 26 dar.) Da die Konjunktion von (Refl), (Trans) und (Sym) das Definiens einer *Äquivalenzrelation* ist, können wir auch einfach sagen: **S5** definiert die Klasse der Äquivalenzrahmen.

Wir möchten, daß die Tafel auf p. 121 Richtigkeits- und Vollständigkeitsresultate über Erweiterungen von **K** anzeigt. Dazu fehlt jetzt noch der Nachweis, daß die Schemata in der Tafel kanonisch sind.

Satz 27. *Alle Schemata in der Tafel auf p. 121 sind kanonisch.*

Beweis. Für jedes der Korrespondenzpaare bestehend aus einem Schema S und einer Bedingung (Bed) zeigen wir, daß der kanonische Rahmen $\mathcal{K}^{\mathbf{L}}$ (**L** eine beliebige Logik, die S enthält) die Eigenschaft (Bed) hat.

Universal quantifizierte Bedingungen, wie Transitivität oder Symmetrie, sind recht einfach zu verifizieren. Bedingungen mit existentieller Quantifikation, wie Erweiterbarkeit, Dichte oder Konvergenz, stellen dagegen eine etwas größere Herausforderung dar. Wir geben hier jeweils zwei Beispiele.

Es sei zunächst an die Definition der Relation R im kanonischen Rahmen erinnert:

I. Rab gdw $\forall A : \Box A \in a \;\Rightarrow\; A \in b$ (d.h. $a^{\Box} \subseteq b$),

wobei a und b hier maximal konsistente Mengen im Sinne von **L** sind. Es ist leicht zu sehen (M3!), daß auch folgendes für solche Mengen a und b gilt:

II. Rab gdw $\forall A$: $A \in b$, dann $\Diamond A \in a$.

Erinnert sei auch daran, daß jeder Punkt im kanonischen Rahmen für **L** mit S alle Instanzen von S enthält (M6!) und daß die Punkte unter Modus Ponens abgeschlossen sind.

Ad 4: Wir zeigen, daß R im kanonischen Rahmen für **L** mit 4 transitiv ist.

(1)	Rab	Annahme
(2)	Rbc	Annahme (zz Rac)
(3)	$\Box A \in a$	Annahme (zz $A \in c$, nach I)
(4)	$\Box A \to \Box\Box A \in a$	Schema 4
(5)	$\Box\Box A \in a$	3,4
(6)	$\Box A \in b$	1, nach I
(7)	$A \in c$	2,6, nach I

Ad 5: Wir zeigen, daß R im kanonischen Rahmen für **L** mit 5 euklidisch ist.

(1)	Rab	Annahme
(2)	Rac	Annahme (zz Rbc)
(3)	$\Box A \in b$	Annahme (zz $A \in c$, nach I)
(4)	$\Diamond\Box A \in a$	1,3, nach II
(5)	$\Diamond\Box A \to \Box A \in a$	Schema 5
(6)	$\Box A \in a$	4,5
(7)	$A \in c$	2,6, nach I

Ad 4_c: Wir zeigen, daß R im kanonischen Rahmen für **L** mit 4_c dicht ist. Wir nehmen also an, daß Rab und zeigen, daß es einen Punkt c gibt mit Rac und Rcb. D.h., wir suchen einen Punkt c mit $a^\Box \subseteq c$ und $c^\Box \subseteq b$. Wir betrachten die Formelmenge

$$C = a^\Box \cup \{\Diamond A : A \in b\}.$$

Wenn C konsistent ist, dann gibt es nach Lindenbaum eine Erweiterung $c \supseteq C$ im kanonischen Rahmen. Wir wissen, daß $a^\Box \subseteq C \subseteq c$, also Rac (nach I). Ferner gilt: $\forall A : A \in b \Rightarrow \Diamond A \in C$ $(\subseteq c)$; also auch Rcb (nach II). Die Menge c erfüllt so die gewünschte Bedingung.

Wir zeigen nun, daß C konsistent ist. Angenommen, das ist nicht der Fall. Dann gibt es Formeln (1) $A_1, ..., A_m \in a^\Box$ und (2) $\Diamond B_1, ..., \Diamond B_n \in \{\Diamond B : B \in b\}$, welche zusammen **L**-inkonsistent sind. Wir wollen die Konjunktionen dieser Formeln jeweils zu A und $\Diamond B$ abkürzen. Die letzte Abkürzung ist gerechtfertigt, da **L** regulär ist, d.h. $\Diamond A \wedge \Diamond A' \leftrightarrow \Diamond(A \wedge A')$ (und ebenso für $\Box$). Aus der Annahme, daß C inkonsistent ist, folgt so, daß (3) $A \to \neg\Diamond B$ in **L**. Aus (1) und (2) folgt jeweils (unter Verwendung der Abkürzung)

$$(4)\ \Box A \in a \quad \text{und} \quad (5)\ B \in b.$$

Auf (3) wenden wir RM an und erhalten $\Box A \to \Box\neg\Diamond B$ in **L**, woraufhin aus (4) folgt, daß $\Box\neg\Diamond B \in a$, d.h. (5) $\Box\Box\neg B \in a$. Jetzt verwenden wir das Schema 4_c, um aus (5) auf (6) $\Box\neg B \in a$ zu schließen. Wiederum aufgrund der Annahme Rab folgt aus (6), daß $\neg B \in b$. Aber nach (5) wäre b nun inkonsistent, d.h. kein Punkt im kanonischen Rahmen – Widerspruch!

Ad G: Wir beweisen, daß R konvergent ist im kanonischen Rahmen für **L** mit G. Dazu nehmen wir an, daß Rab und Rac und zeigen, daß es dann einen Punkt d gibt mit Rbd and Rcd. Wir beginnen mit der Formelmenge

$$D = b^\Box \cup c^\Box.$$

Es ist klar, daß RbD und RcD. Wenn D konsistent ist, dann garantiert das Lindenbaum-Lemma eine maximal konsistente Erweiterung d von D mit den gewünschten Eigenschaften Rbd und Rcd.

Angenommen, D ist nicht konsistent. Dann gibt es Formeln $B_1, ..., B_m \in b^\Box$ und $C_1, ..., C_n \in c^\Box$ so, daß deren Konjunktion **L**-inkonsistent ist. Die Konjunktion von $B_1, ..., B_m$ wollen wir zu B, die Konjunktion von $C_1, ..., C_n$ zu C abkürzen. Es folgt unmittelbar, daß

$$(1)\ \Box B_1, ..., \Box B_m \in b \quad \text{und} \quad (2)\ \Box C_1, ..., \Box C_n \in c.$$

Da **L** regulär ist, folgen aus (1) und (2) jeweils

$$(3)\ \Box B \in b \quad \text{und} \quad (4)\ \Box C \in c.$$

Die Annahme der Inkonsistenz von D bedeutet, daß $B \to \neg C$ in **L**. Daraus folgt aufgrund von RM$\Diamond$, $\Diamond B \to \Diamond\neg C$ und also (5) $\Diamond B \to \neg\Box C$ in **L**. Aus Rab und (3) folgt (wie eingangs bemerkt), daß $\Diamond\Box B \in a$, woraus wir per Schema G auf $\Box\Diamond B \in a$ schließen dürfen. Daraus folgt aufgrund der weiteren Annahme Rac, daß $\Diamond B \in c$. Aber jetzt können wir mit (5) auf $\neg\Box C \in c$ schließen. Somit wäre nach (4) c inkonsistent und also, entgegen unserer Annahme, kein Punkt im kanonischen Rahmen. ∎

Aus diesem Satz zusammen mit dem Korrespondenzsatz 25 folgt nun unmittelbar der Determinationssatz:

SATZ 28. (RICHTIGKEIT & VOLLSTÄNDIGKEIT) *Für alle Paare (Bed,S) in der Tafel auf p. 121 gilt:* $Th(Bed(\mathbb{K})) = \mathbf{K} + S$. *Ferner gilt für jede Kombination von korrespondierenden Bedingungen und Schemata:*

$$Th(Bed_1(\mathbb{K}) \cap \cdots \cap Bed_n(\mathbb{K})) = \mathbf{K} + S_1 + \cdots S_n.$$

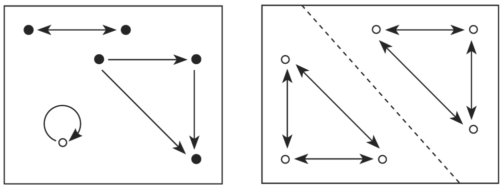
Insbesondere haben wir das folgende

KOROLLAR 29.
1. *Die Klasse der reflexiven und transitiven Rahmen determiniert das System* **S4** (= **KT4**).
2. *Die Klasse der reflexiven und euklidischen Rahmen (d.h. die Klasse der Äquivalenzrahmen) determiniert das System* **S5** (= **KT5**).

Damit sind wir wieder bei einem alten Bekannten: der Logik logischer Notwendigkeit, **S5**.

9. Die Vollständigkeit von S5 für universale Rahmen

Im Bild unten sind im linken Rahmen Punkte zu sehen, die symmetrisch, transitiv, und reflexiv zueinander stehen. (Ungefüllte Punkte sollen reflexive Punkte darstellen, d.h. solche Punkte a mit Raa.)



Im rechten Rahmen sind die links einzeln dargestellten Eigenschaften kombiniert: Die Relation ist jetzt durchweg reflexiv, transitiv und symmetrisch. also eine Äquivalenzrelation.[18] Zugleich zeigt der rechte Rahmen eine verdoppelte Punktestruktur. Man sieht, daß der Bereich in zwei Teilbereiche zerfällt (durch die gestrichelte Linie angedeutet), in denen jeder Punkt zu sich selbst und zu allen anderen in der Relation steht. D.h. innerhalb dieser Teilbereiche – aber nicht im gesamten Bereich – ist die Relation R *universal.*

Offenbar sind alle Äquivalenzrahmen so wie im Bild rechts *pars pro toto* dargestellt: Sie zerfallen vollständig in disjunkte Äquivalenzklassen, innerhalb derer die Relation universal ist. Nach dem Korollar 29.2 zum Determinationssatz wissen wir, daß **S5** richtig und vollständig ist im Hinblick auf genau solche Rahmen. Aber um $\Box$ als *logische* Notwendigkeit zu interpretieren, hätten wir gern die Vollständigkeit von **S5** in Bezug auf Rahmen, die so sind wie die Teilrahmen im rechten Rahmen, d.h. Rahmen, in denen R universal ist. Erst dann gilt:

> $\Box A$ ist an einem Punkt genau dann wahr, wenn A an *allen* Punkten wahr ist.

Wenn **S5** die Logik logischer Notwendigkeit sein soll, dann brauchen wir daher dieses Resultat:

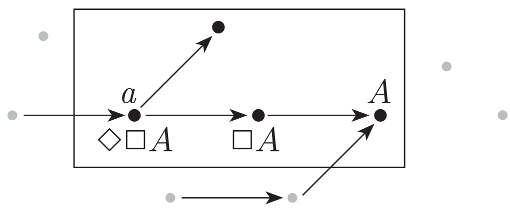
SATZ 30. *Das System* **S5** *ist richtig und vollständig für alle Rahmen* (W, R) *in denen* R **universal** *in* W *ist.*

Der Nachweis der *Richtigkeit* ist trivial: Wenn wir A in einem Modell verifizieren können unter der Annahme, daß R eine Äquivalenzrelation ist, dann gelingt das sicher auch unter der stärkeren Annahme, daß R universal ist. Für den Nachweis der *Vollständigkeit* würde es reichen, wenn wir zeigen

[18] Zur Terminologie und einigen elementaren Beobachtungen über Relationen siehe Kap. I.2.

könnten, daß es für die Gültigkeit einer Formel gewissermaßen keinen Unterschied macht, ob wir den ganzen Rahmen im rechten Bild betrachten (R reflexiv, symmetrisch und transitiv) oder einen der Teilrahmen (mit R universal). Denn dann könnte eine Formel nicht in einem Äquivalenzrahmen gültig sein ohne zugleich in einem der universalen Teilrahmen gültig zu sein. Diese Überlegung wollen wir im folgenden mit der erforderlichen Genauigkeit umsetzen.

Erzeugte Teilmodelle. Gegeben sei ein Modell $\mathcal{M} = (W, R, I)$. Wir fragen, ob an einem bestimmten Punkt a eine bestimmte Formel A wahr ist. Wenn A Kisten enthält, dann wird uns die Relation R an weitere Punkte verweisen. Für die Beurteilung von A am Punkt a sind aber letztlich nur Punkte relevant, die zu a in der R-Relation stehen. D.h., um zu entscheiden, ob $a \models A$, genügt es ein Teilmodell von $\mathcal{M}$ zu betrachten, nämlich dasjenige Teilmodell, in dem nur diejenigen Punkte vorkommen, an die wir im Laufe der Beurteilung von $a \models A$ verwiesen werden. Ein Diagramm macht das deutlich:



Um am Punkt a die Formel $\Diamond\Box A$ zu beurteilen, müssen wir nicht alle Punkte im Modell betrachten, sondern nur die Punkte im Rechteck, auf welche die Pfeile (d.h. die Relation R) von a ausgehend verweisen. Man beachte:

· Auch wenn die Relation nicht reflexiv ist, gehört der Punkt a, an dem wir die Formel beurteilen möchten, zu den für die Beurteilung relevanten Punkten.
· Und auch wenn die Relation selbst nicht transitiv sein sollte, ist das Verweisen über die Relation auf relevante Punkte eine transitive Relation.

Um alle von a aus relevanten Punkte einzusammeln, bilden wir deshalb den *reflexiven und transitiven Abschluß* R^+ der Relation R:

$$R^0ab \text{ gdw } a = b;$$
$$R^{n+1}ab \text{ gdw } \exists c: \ R^nac \ \& \ Rcb;$$
$$R^+ab \text{ gdw } \exists n \geq 0: \ R^nab.$$

Das vom Punkt a aus erzeugte Teilmodell besteht genau aus den Punkten, die von a aus per R^+ erreichbar sind. Das ist der Inhalt der folgenden

DEFINITION 31. Gegeben ein Modell $\mathcal{M} = (W, R, I)$ mit $a \in W$, ist $\mathcal{M}^a = (W^a, R^a, v^a)$ das *aus a erzeugte Teilmodell von* $\mathcal{M}$, wenn
1. $W^a = \{b \in W : R^+ab\}$;
2. $R^a = R \cap (W^a \times W^a)$, d.h. R^abc gdw Rbc und $b, c \in W^a$;
3. $\forall b \in W^a : \; I^a(P, b) = I(P, b)$.

Der Rahmen (W^a, R^a) ist der aus *a erzeugte Teilrahmen* von (W, R).

Der nächste Satz besagt, daß das aus einem Punkt erzeugte Teilmodell der einzige Teil des Modells ist, der für die Wahrheit einer Formel an diesem Punkt relevant ist.

SATZ 32. (Erzeugte Teilmodelle) *Für jeden Punkt $b \in W^a$ eines aus a erzeugten Teilmodells M^a von $M = (W, R, I)$ gilt (für beliebige Formeln A):*

$$b \models A \text{ in } M^a \text{ gdw } b \models A \text{ in } M.$$

BEWEIS. Um sich von der Wahrheit des Satzes zu überzeugen, genügt ein Blick auf das oben gezeigte Diagramm. Der eigentliche Beweis verfährt per Induktion über den Aufbau von A. (Wir schreiben kurz $b \models^a A$ für '$b \models A$ in M^a'.) Der Fall $A \in$ ATM ist durch die Klausel 3 der Definition gegeben.

Fälle $A = \neg B$ und $A = B \to C$. Hier finden keine Verweise *per* R bzw. R^a statt.

Fall $A = \Box B$, d.h. $b \models^a \Box B \;\Leftrightarrow\; b \models \Box B$.

LR: Wir nehmen an, daß $b \models^a \Box B$; zz $b \models \Box B$, d.h. wenn Rbc, dann $c \models B$ (c beliebig). Die Induktionsannahme (IA) ist: Wenn $x \models^a B$, dann $x \models B$ ($\forall x \in W$).

(1)	$b \models^a \Box B$	Annahme
(2)	Rbc	Annahme (zz $c \models B$)
(3)	$b \in W^a$, d.h. R^+ab	1, Def. W^a
(4)	R^+bc, d.h. $c \in W^a$	2,3, Def. R^+, Def. W^a
(5)	$\forall x : R^abx \Rightarrow x \models^a B$	1
(6)	R^abc	2,3,4 Def. R^a
(7)	$c \models^a B$	5,6
(8)	$c \models B$	7, IA

RL: IA: Wenn $x \models B$, dann $x^a \models B$ ($\forall x \in W$).

(1)	$b \models \Box B$	Annahme
(2)	R^abc	Annahme (zz $c \models^a B$)
(3)	$\forall x \in W : Rbx \Rightarrow x \models B$	1
(4)	Rbc	2, Def. R^a
(5)	$c \models B$	3,4
(6)	$c \models^a B$	5, IA

∎

Zurück zu S5. Der Satz über erzeugte Teilmodelle verhilft uns nun schnell zum Beweis des Satzes 30. Dazu betrachten wir das kanonische Modell $\mathcal{M} = (W, R, I)$ für **S5** und eine beliebige Formel A, die in **S5** nicht ableitbar ist. Dann gibt es einen Punkt a in $\mathcal{M}$ mit $a \not\models A$. Jetzt erzeugen wir aus $\mathcal{M}$ das Teilmodell $\mathcal{M}^a$.

1. $a \not\models^a A$ aufgrund des Satzes über erzeugte Teilmodelle.
2. In $\mathcal{M}^a$ ist R^a universal. Denn da R reflexiv und transitiv ist, so ist $R = R^+$ und also $W^a = R(a)$. Da R außerdem noch symmetrisch ist, so ist R eine Äquivalenzrelation; also ist W^a eine Äquivalenzklasse unter R. Daraus folgt aber, daß R universal ist in W^a, und da R^a nur die Einschränkung von R auf Elemente in W^a ist, so ist auch R^a universal in W^a.

Für jedes Nicht-Theorem A von **S5** gibt es so ein universales Modell – basierend auf einer Äquivalenzklasse des kanonischen Modells für **S5** –, welches A falsifiziert. Kontraponiert: Wenn $A \in$ **S5**, dann ist A wahr in allen universalen Modellen, d.h. gültig in der Klasse $Univ(\mathbb{K})$ der universalen Rahmen.

10. Die Grenzen der Methode

Undefinierbare Rahmenklassen. Wir haben gesagt (Def. 19): Eine Logik **L** definiert eine Rahmenklasse $\mathbb{L}$ genau dann, wenn $Rm(\mathbf{L}) = \mathbb{L}$. Dementsprechend ist eine Rahmenklasse $\mathbb{R}$ modal *definierbar*, falls es eine Logik **L** gibt so, daß $Rm(\mathbf{L})$ genau $\mathbb{R}$ ist; ferner ist eine Rahmenbedingung (Bed) modal definierbar, falls es ein Formelschema S gibt so, daß Bedingung und Schema korrespondieren, d.h. $Bed(\mathbb{K}) = Rm(S)$. Nun fragen wir:

· Ist jede Rahmenklasse modal definierbar? Äquivalent:
· Gibt es für jede Rahmenbedingung ein sie definierendes, d.h. ein korrespondierendes Schema?

Einen ersten Hinweis darauf, daß das nicht so ist, gibt der folgende Satz.

Satz 33. *Wenn $\mathbb{R}_1 \neq \mathbb{R}_2$ und $Th(\mathbb{R}_1) = Th(\mathbb{R}_2)$, dann ist $\mathbb{R}_1$ oder $\mathbb{R}_2$ nicht modal definierbar.*

Beweis. Zum Antezedens des Satzes nehmen wir – für *reductio* – die weitere Annahme hinzu, daß es Logiken $\mathbf{L}_1$ und $\mathbf{L}_2$ gibt mit

$$(*) \qquad \mathbb{R}_1 = Rm(\mathbf{L}_1) \quad \text{und} \quad \mathbb{R}_2 = Rm(\mathbf{L}_2).$$

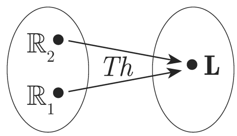
Aus $\mathbb{R}_1 \neq \mathbb{R}_2$ und $(*)$ folgt, daß (1) $\mathbf{L}_1 \neq \mathbf{L}_2$. Nun gilt aber allgemein, daß

$$Th(Rm(\mathbf{L}_1)) = \mathbf{L}_1 \quad \text{und} \quad Th(Rm(\mathbf{L}_2)) = \mathbf{L}_2.$$

Somit folgt aus (1), daß $Th(Rm(\mathbf{L}_1)) \neq Th(Rm(\mathbf{L}_2))$. Nach Einsetzen aufgrund von $(*)$ erhalten wir so einen Widerspruch zu unserer Annahme, daß $Th(\mathbb{R}_1) = Th(\mathbb{R}_2)$. ■

Korollar 34. *Wenn für nicht äquivalente Bedingungen* (Bed$_1$) *und* (Bed$_2$)*,* $Th(Bed_1(\mathbb{K})) = Th(Bed_2(\mathbb{K})$*, und* (Bed$_1$) *modal definierbar ist, dann ist* (Bed$_2$) *nicht modal definierbar.*

Das Antezedens des gerade bewiesenen Satzes faßt die Annahme ins Auge, daß die Funktion Th nicht injektiv ist. Im Bild:



Genau diese Situation belegt das Beispiel von **S5**. Hier haben wir zwei Rahmenklassen, $Univ(\mathbb{K})$ und $Aeq(\mathbb{K})$, mit identischer Theorie, nämlich **S5**, und also dürfen wir schließen, daß eine der beiden Rahmenklassen modal nicht definierbar ist. Nun wissen wir (aufgrund von Satz 25), daß $Aeq(\mathbb{K})$ definierbar ist, denn

$$Aeq(\mathbb{K}) = Rm(\mathbf{S5}).$$

Also ist es die Klasse der universalen Rahmen, welche nicht definierbar ist.

Weitere nicht definierbare Rahmenklassen sind:

1. die Klasse der irreflexiven Rahmen: $(\forall a)$ nicht Raa,
2. die Klasse der asymmetrischen Rahmen: $(\forall ab)$ wenn Rab, dann nicht Rba.
3. die Klasse der antisymmetrischen Rahmen: $(\forall abc)$ wenn Rab und Rbc, dann $b = c$.[19]

Auch in diesen Fällen erhalten wir, wie für **S5**, multiple Vollständigkeitsresultate und können so Satz 33 anwenden. So stimmen zum Beispiel die Theorie der transitiven Rahmen, die Theorie der transitiven und irreflexiven Rahmen, als auch die Theorie der endlichen transitiven Rahmen überein: Sie alle koinzidieren mit dem System **K4**. Aber $Trans(\mathbb{K})$ ist ebensowenig identisch mit $Trans(\mathbb{K}) \cap Irr(\mathbb{K})$ wie mit $Trans(\mathbb{K}) \cap Endl(\mathbb{K})$. Nach Satz 33

[19] Die modale Undefinierbarkeit der Irreflexivität haben wir schon im Kapitel II.9 gezeigt. Für den Nachweis weiterer Undefinierbarkeiten siehe z.B. [31].

schließen wir daher, daß Transitivität oder Irreflexivität bzw. Transitivität oder Endlichkeit nicht definierbar sind. Aber wir wissen (aufgrund von Satz 25), daß das Schema 4 mit der Bedingung der Transitivität korrespondiert und also **K4** die transitiven Rahmen definiert. Also sind die Klassen $Irr(\mathbb{K})$ und $Endl(\mathbb{K})$ nicht definierbar.

Multiple Determinationsresultate, wie im Falle von **S5** im Hinblick auf universale als auch auf Äquivalenzrahmen, deuten eine ganz allgemeine Strategie an, wie sich die Undefinierbarkeit bestimmter Rahmenklassen zeigen läßt. Wir wollen hier den speziellen Sinn von Definierbarkeit im Sinne von Korrespondenz betrachten. (Die Verallgemeinerung liegt dann auf der Hand.) So ist eine (konsistente) Rahmenbedingung (Bed) (per Korrespondenz) nicht definierbar, wenn es kein Formelschema S gibt so, daß S mit (Bed) korrespondiert. Um zu zeigen, daß (Bed) in diesem Sinne nicht definierbar ist, könnten wir per *reductio* so verfahren:

1. Wir nehmen an, es gäbe ein zu (Bed) korrespondierendes Schema S: Also (1) $\forall\mathcal{R} : \mathcal{R} \models S$ gdw $Bed(R)$.
2. Wir betrachten nun einen Rahmen $\mathcal{R}$, welcher die Bedingung erfüllt. Nach (1) schließen wir auf (2) $\mathcal{R} \models S$.
3. Jetzt verschaffen wir uns eine Methode, die aus dem (Bed)-Rahmen $\mathcal{R}$ einen (3) ($\neg$Bed)-Rahmen $\mathcal{R}'$ konstruiert so, daß gleichwohl (4) $Th(\mathcal{R}) \subseteq Th(\mathcal{R}')$.
4. Aus (2) und (4) folgt, daß $\mathcal{R}' \models S$ und also aus (1), daß $Bed(\mathcal{R}')$ – im Widerspruch zu (3).

Die Durchführbarkeit der skizzierten Strategie hängt offenbar davon ab, daß wir gegebene Rahmen in neue Rahmen überführen können, wobei einerseits bestimmte Eigenschaften verlorengehen, andererseits aber die Menge der gültigen Formeln bewahrt bleibt.

Eine besonders einfache Art aus einem Rahmen einen neuen Rahmen zu konstruieren und dabei die gültigen Formeln zu bewahren, ist die Konstruktion eines isomorphen Bildes eines gegebenen Rahmen. Zwei Rahmen (W, R) und (W', R') sind *isomorph*, wenn es eine Bijektion $f : W \longrightarrow W'$ gibt und $R'f(a)f(b)$ gdw Rab. Es ist leicht zu sehen, daß in diesem Fall $Th(W, R) = Th(W', R')$, die Bewahrung von Gültigkeit also in beide Richtungen geht. Zugleich ist aber auch klar, daß sich isomorphe Rahmen im Hinblick auf die hier interessierenden Eigenschaften nicht unterscheiden. Im folgenden betrachten wir drei interessantere Konstruktionen, um die Undefinierbarkeit bestimmter Eigenschaften nachzuweisen.

DEFINITION 35.

1. Ein *p-Morphismus* von $\mathcal{R} = (W, R)$ nach $\mathcal{R}' = (W', R')$ ist eine Funktion $f : W \longrightarrow W'$, welche diese Bedingungen erfüllt:

(pM1) $(\forall ab)$ Wenn Rab, dann $R'f(a)f(b)$; und
(pM2) $(\forall ab')$ wenn $R'f(a)b'$, dann $\exists b$: $f(b) = b'$ und Rab.

2. $\mathcal{R}'$ ist ein *p-morphes Bild* von $\mathcal{R}$, wenn es einen surjektiven p-Morphismus $f : W \longrightarrow W'$ gibt.

SATZ 36. (p-Morphismus) *Es sei f ein p-Morphismus von $\mathcal{R} = (W, R)$ nach $\mathcal{R}' = (W', R')$ und I, I' seien Interpretation auf $\mathcal{R}$ bzw. $\mathcal{R}'$ so, daß*

(Atm) $I(P, a) = I'(p, f(a))$, für alle Atome P und Punkte $a \in W$.

Dann $\forall A \in FML$ und $a \in W$,

$$a \models A \textit{ in } (\mathcal{R}, I) \textit{ gdw } f(a) \models A \textit{ in } (\mathcal{R}', I').$$

BEWEIS. Induktion über den Aufbau von A. Die Basis ist durch (Atm) gegeben. Die Fälle $A = \neg B$ und $A = B \wedge C$ folgen unmittelbar aus der Induktionsannahme (IA).

Ad $A = \Box B$: In der LR-Richtung nehmen wir an, daß (1) $\forall x : Rax \Rightarrow x \models B$ und (2) $R'f(a)b'$; zz $b' \models B$. Dann gibt es nach (pM2) ein $b \in W$ mit (3) $b' = f(b)$ und (4)Rab. Es folgt aus (1) und (2), daß (5) $b \models B$. Auf (5) können wir die IA anwenden und erhalten $f(b) \models B$, d.h. nach (3), $b' \models B$.

In umgekehrter Richtung gehen wir aus von (1) $\forall x : R'f(a)x \Rightarrow x \models B$ und (2) Rab; zz $b \models B$. Nach (pM1) folgt nun aus (2), daß (3) $R'f(a)f(b)$. Also folgt aus (1) weiter, daß (4) $f(b) \models B$. Auf (4) wenden wir die IA an und erhalten so $b \models B$ ∎

DEFINITION 37. Es seien $\mathcal{R}_1 = (W_1, R_1)$ und $\mathcal{R}_2 = (W_2, R_2)$ so, daß $W_1 \cap W_2 = \emptyset$. Dann ist der Rahmen

$$\mathcal{R}_1 \oplus \mathcal{R}_2 = (W_1 \cup W_2, R_1 \cup R_2)$$

die *disjunkte Vereinigung* der *Disjunkte* $\mathcal{R}_1$ und $\mathcal{R}_2$.[20]

SATZ 38. (Disjunkte Vereinigung) *Es sei $\mathcal{R}'$ die disjunkte Vereinigung von $\mathcal{R}_1$ und $\mathcal{R}_2$ und $I'(P, a) = 1$ gdw $I_1(P, a) = 1$ oder $I_2(P, a) = 1$* ($\forall P \in$ *ATM*). *Dann $\forall A \in FML$ und $a \in W_i$* $(i \in \{1, 2\})$,

$$a \models A \textit{ in } (\mathcal{R}_i, I_i) \textit{ gdw } a \models A \textit{ in } (\mathcal{R}', I').$$

[20] Statt, wie hier, nur zwei disjunkte Rahmen zu betrachten, können wir die Definition auch natürlich verallgemeinern zur disjunkten Vereinigung von Rahmenklassen mit mehr als zwei Elementen.

Beweis. Induktion über den Aufbau von A. Wir gehen aus von einem beliebig gewählten Punkt a in W_i.

Fall A ist ein Atom P: Die gewünschte Gleichung $I_i(P,a) = I'(P,a)$ folgt aus der Definition von I'.

Fall A ist eine Negation oder eine Konjunktion: Folgt jeweils unmittelbar aus der IA.

Fall $A = \Box B$: (Es sei $a \models_i A$ eine Abkürzung für $a \models A$ in $(\mathcal{R}_i, I_i)$.)

$\Rightarrow$: Wir nehmen an, daß (1) $a \models_i \Box B$ und (2) $R'ab$; zz $b \models' B$. Aus (1) folgt (3) $\forall x : R_i ax \Rightarrow x \models_1 B$. Aus (2) folgt, daß $R_i ab$ oder $R_j ab$ $(i \neq j)$. Aber da $a \in W_1$, so nicht $R_j ab$; also (4) $R_i ab$. So folgt aus (3), daß $b \models_1 B$, worauf wir die IA zur Konklusion $b \models' B$ anwenden.

$\Leftarrow$: Folgt aus $R_i \subseteq R'$. ■

Schließlich sei an die Definition 31 (p. 129) erzeugter Teilmodelle und -rahmen, sowie an den zugehörigen Satz erinnert:

Satz 32. (Erzeugte Teilmodelle) *Es sei* $(\mathcal{R}^a, I^a)$ *ein aus* a *erzeugtes Teilmodell von* $(\mathcal{R}, I)$ *(also* $\mathcal{R}^a$ *ein erzeugter Teilrahmen von* $\mathcal{R}$*). Dann* $\forall A \in FML$ *und* $b \in W^a$,

$$b \models A \text{ in } (\mathcal{R}^a, I^a) \text{ gdw } b \models A \text{ in } (\mathcal{R}, I).$$

Korollar 39. (Aus den drei vorangegangenen Sätzen.)

1. *Wenn* $\mathcal{R}'$ *ein p-morphes Bild von* $\mathcal{R}$ *ist, dann* $Th(\mathcal{R}) \subseteq Th(\mathcal{R}')$.
2. *Wenn* $\mathcal{R}'$ *ein erzeugter Teilrahmen von* $\mathcal{R}$ *ist, dann* $Th(\mathcal{R}) \subseteq Th(\mathcal{R}')$.
3. *Wenn* $\mathcal{R}_1$ *und* $\mathcal{R}_2$ *die Disjunkte von* $\mathcal{R}'$ *sind, dann* $Th(\mathcal{R}_1) \cap Th(\mathcal{R}_2) \subseteq Th(\mathcal{R}')$.
4. *(a) Wenn in 1 und 2* $Bed(\mathcal{R})$ *und* $\neg Bed(\mathcal{R}')$*, dann ist (Bed) nicht definierbar; (b) wenn in 3* $Bed(\mathcal{R}_1)$, $Bed(\mathcal{R}_2)$ *und* $\neg Bed(\mathcal{R}')$*, dann ist (Bed) nicht definierbar.*

Ad 1: Angenommen, das sei nicht so. Dann gibt es ein Modell $(\mathcal{R}', I')$ mit einem Punkt a' so, daß $a' \not\models A$, für irgendeine Formel A mit $(*)$ $\mathcal{R} \models A$. Wenn f der angenommene p-Morphismus von $\mathcal{R}$ nach $\mathcal{R}'$ ist, dann sei I auf $\mathcal{R}$ so, daß $I(P,a) = I'(P, f^{-1}(a'))$. (Da f surjektiv ist, so ist die Umkehrung f^{-1} wohldefiniert.) Nach Satz 36 folgt nun, daß $f^{-1}(a') \not\models A$ in $(\mathcal{R}, I)$, d.h. $\mathcal{R} \not\models A$ – im Widerspruch zu $(*)$.

Ad 2: Für *reductio* nehmen wir an, daß $(*)$ $\mathcal{R} \models A$, während es die Einschränkung I' auf W' einer Interpretation I auf W und einen Punkt $a \in W'$ gibt mit $a \not\models A$ in $(\mathcal{R}', I')$. Aber dann $a \not\models A$ in $(\mathcal{R}, I)$ nach Satz 32, d.h. $\mathcal{R} \not\models A$ – Widerspruch zu $(*)$.

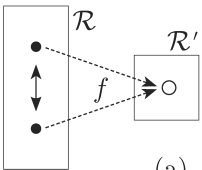
Ad 3: Wir nehmen wieder für *reductio* an, daß $(*)$ $\mathcal{R}_1 \models A$ und $\mathcal{R}_2 \models A$, während $\mathcal{R}' \not\models A$. Dann gibt es ein Modell $(\mathcal{R}', I')$ mit einem Punkt a so, daß $a \not\models' A$. Aber $a \in W_1 \cup W_2$. Also muß es ein Modell auf $\mathcal{R}_1$ oder auf $\mathcal{R}_2$ geben so, daß nach Satz 38 $a \not\models_1 A$ oder $a \not\models_2 A$ – im Widerspruch zu $(*)$.

Ad 4: Siehe die Beschreibung unserer Strategie oben (p. 132). ∎

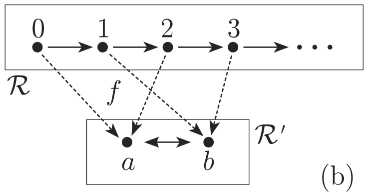
Für disjunkte Vereinigungen halten wir eine einfach zu beweisende Verallgemeinerung des Korollars fest:

Wenn $M = \{R_i : i \in I\}$ ($I \subseteq \mathbf{N}$ eine Indexmenge) eine Menge paarweise disjunkter Rahmen ist, dann ist $\bigcap\{Th(R_i) : i \in I\} \subseteq Th(R')$. — Wenn $Bed(R_i)$ für alle $i \in I$ und $\neg Bed(R')$, dann ist (Bed) nicht definierbar.

Wir wollen das Korollar anhand einiger Beispiele illustrieren. In den Diagrammen (a) und (b) sehen wir Beispiele von p-Morphismen f.



(a)



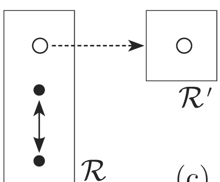
(b)

Daß f in (a) ein p-Morphismus ist, ist unschwer zu verifizieren. In (b) werden die natürlichen Zahlen auf zwei Punkte a und b abgebildet. Die Vorschrift für f lautet hier:

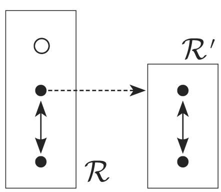
$$f(n) = \begin{cases} a, & \text{wenn } n \text{ gerade ist;} \\ b & \text{anderenfalls.} \end{cases}$$

Daß nun f die Bedingung (pM2) erfüllt, ist sicher nicht offensichtlich; siehe den Beweis in [31, p. 60]. In Diagramm (a) ist in $\mathcal{R}$ die Relation *irreflexiv* (und daher auch *nicht reflexiv*). Diese Eigenschaft wird beim Übergang zum p-morphen Bild $\mathcal{R}'$ nicht bewahrt. Also ist Irreflexivität (wie die Nicht-Reflexivität) nicht definierbar. Im Diagramm (b) ist die (Nachfolger-) Relation *nicht symmetrisch*. Im p-morphen Bild $\mathcal{R}'$ ist die Relation symmetrisch. Also ist Nicht-Symmetrie nicht definierbar.

Die folgenden zwei Diagramme zeigen die Erzeugung von Teilrahmen.



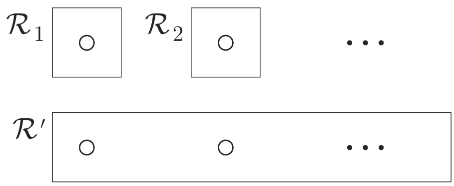
(c)



(d)

Im Diagramm (c) ist die Relation im Rahmen $\mathcal{R}$ *nicht universal* – im Gegensatz zur Relation in $\mathcal{R}'$. Also ist Nicht-Universalität nicht definierbar. Ebenso zeigt das Beispiel in (c) erneut die Undefinierbarkeit der *Nicht-Reflexivität.* In (d) hat die Relation die Eigenschaft *nicht irreflexiv* zu sein – eine Eigenschaft, die nach $\mathcal{R}'$ nicht übertragen wird. Ebenso wird die Eigenschaft *nicht symmetrisch* zu sein auf den generierten Teilrahmen übertragen.

Schließlich kommen wir zur Vereinigung disjunkter Rahmen $\mathcal{R}_1, \mathcal{R}_2, ...$ zu einem Rahmen $\mathcal{R}'$ in folgendem Beispiel:



(e)

Hier stellen wir zunächst fest (wobei es auf die angedeutete Reflexivität der Punkte nicht ankommt), daß alle $\mathcal{R}_i$ *endlich* sind, während $\mathcal{R}'$ es nicht ist. Endlichkeit ist also keine definierbare Eigenschaft von Rahmen. Dasselbe gilt offensichtlich auch für beliebige Mächtigkeitsaussagen "der Rahmen hat höchstens n Elemente". In umgekehrter Richtung sind alle R_i aus (Punkten in) R' generiert. Diese Umkehrung führt zu komplementären Beobachtungen: Unendlichkeit, sowie Mächtigkeitsaussagen der Form "ein Rahmen hat mindestens n Punkte" sind nicht definierbar. Ferner sind im Gegensatz zu $\mathcal{R}'$ alle $\mathcal{R}_i$ *universal*, womit noch einmal die Undefinierbarkeit der Universalität gezeigt ist.

Wir können demnach festhalten, daß es für die Definierbarkeit einer Rahmeneigenschaft notwendig ist, daß diese Eigenschaft übertragen wird auf (1) p-morphe Bilder, (2) generierte Teilrahmen, sowie auf (3) die Vereinigung mit disjunkten Rahmen, die ebenfalls die Eigenschaft aufweisen. Mit (1–3) sind wir tatsächlich schon in der Nähe einer nicht nur notwendigen, sondern auch hinreichenden Bedingung für die Definierbarkeit einer erstufigen Rahmeneigenschaft (Bed). Eine solche Bedingung wird im folgenden Satz identifiziert:

Satz 40. (Goldblatt und Thomason) *Eine Rahmenklasse $\mathbb{R}$ ist genau dann modal definierbar, wenn $\mathbb{R}$ abgeschlossen ist unter der Konstruktion p-morpher Bilder, generierter Teilrahmen, disjunkter Vereinigungen und Ultrafiltererweiterungen.*

Nur die letzte hier genannte Bedingung haben wir hier nicht besprochen, da sie den Rahmen unserer Betrachtung sprengen würde; vgl. dazu z.B. [31].

Bisimulation. Eine wesentliche Eigenart der drei gerade besprochenen Rahmenkonstruktionen ist diese: Für jeden Punkt im Originalrahmen und sein Gegenstück im konstruierten Rahmen ist die Art der R-Vernetzung mit anderen Punkten völlig gleich. Das ist der Grund dafür, warum die Formelinduktion in den Beweisen der Sätze 32, 36 und 38 auch im Falle modaler Formeln durchgeht. Denn die Wahrheit von $\Box A$ am Punkt a hängt allein davon ab, *welche* Punkte a sieht und *was* a dort sieht. Wenn zwischen zwei Modellen Punkte so gepaart werden können, daß sie jeweils den gleichen Ausblick auf qualitativ gleiche Punkte bieten, dann können die Punkte solcher Paare offenbar keine verschiedenen Aussagen darüber machen, "was sie sehen": Was an dem einen Punkt des Paares wahr ist, muß auch an dem anderen Punkt wahr sein.

Ein Punkt und sein p-morphes Bild, ein Punkt a in einem Modell und derselbe Punkt im aus a generierten Teilmodell bzw. in einem durch Vereinigung mit disjunkten Modellen entstandenen größeren Modellen – all dies sind Paarungen der gerade beschriebenen Art. Die allgemeine Art einer solchen Paarung nennen wir eine *Bisimulation.*[21]

DEFINITION 41. 1. Eine (nichtleere) Relation $H \subseteq W \times W'$ ist eine *Bisimulation* zwischen Rahmen $\mathcal{R} = (W, R)$ und $\mathcal{R}' = (W', R')$ gdw $\forall a, a'$:[22]

(Hin) $(\forall b)$ Wenn Haa' und Rab, dann $\exists b'$: $R'a'b'$ und Hbb' ; und
(Her) $(\forall b')$ wenn Haa' und $R'a'b'$, dann $\exists b$: Rab und Hbb'.

2. Eine Bisimulation zwischen Rahmen ist auch eine *zwischen Modellen* $(\mathcal{R}, I)$ und $(\mathcal{R}', I')$ auf diesen Rahmen, wenn darüber hinaus gilt:

(Atm) $(\forall a, a')$ Wenn Haa', dann $I(P, a) = I'(P, a')$ $(\forall P \in \text{ATM})$.

SATZ 42. (Bisimulation) *Es sei H eine Bisimulation zwischen $\mathcal{M}$ und $\mathcal{M}'$. Dann $\forall A \in FML$, a in $\mathcal{M}$ und a' in $\mathcal{M}'$: Wenn Haa', dann*

$$a \models A \text{ gdw } a' \models A.$$

BEWEIS. Induktion über die Komplexität von A. Die Basis ist durch (Atm) gegeben. Die Fälle $A = \neg B$ und $A = B \wedge C$ folgen unmittelbar aus der Induktionsannahme (IA).

[21] Die Verwendung von Bisimulationen in der Modallogik geht auf van Benthem zurück. Zur Geschichte und Anwendungsbreite von Bisimulationen ganz allgemein siehe [255].
[22] Die Variablen seien sortiert: ungestrichene aus W, gestrichene aus W'.

Ad $A = \Box B$: (LR) Wir nehmen an, daß (1) $a \models \Box B$; zz $a' \models \Box B$. Wir nehmen ferner an, daß (2) $R'a'b'$; es bleibt zz, daß $b' \models B$. Aus Haa' und (2) folgt nach (Her), daß es einen Punkt b gibt mit (3) Rab und (4) Hbb'. Aus (1) und (3) folgt nun $b \models B$, und also aus (4) und der IA, $b' \models B$. – (RL) verfährt spiegelbildlich unter Verwendung von (Hin). ∎

Lemma 43.

1. *Ist $\mathcal{M}'$ ein aus $\mathcal{M}$ erzeugtes Teilmodell, dann gibt es eine Bisimulation zwischen $\mathcal{M}$ und $\mathcal{M}'$.*
2. *Ist $\mathcal{M}'$ eine disjunkte Vereinigung mit Disjunkt $\mathcal{M}$, dann gibt es eine Bisimulation zwischen $\mathcal{M}$ und $\mathcal{M}'$.*
3. *Ist $\mathcal{M}'$ eine p-morphes Bild von $\mathcal{M}$, dann gibt es eine Bisimulation zwischen $\mathcal{M}$ und $\mathcal{M}'$.*

Beweis. Für 1 und 2 sei H die Identität in W', $\{(a, a) : a \in W'\}$; für 3 sei Haa' gdw $f(a) = a'$. ∎

Die Sätze 32, 38 und 36 folgen nun jeweils als Korollar aus dem Bisimulationssatz 42.

Man beachte, daß aus dem Bisimulationssatz kein Korollar wie 39 folgt, welches Undefinierbarkeitsresultate im bisherigen Sinne, d.h. im Hinblick auf Rahmenklassen herbeiführen könnte. Um das einzusehen, betrachten wir nochmals das Diagramm (a) auf p. 135. Der Leser möge sich davon überzeugen, daß sowohl die Funktion $f : \mathcal{R} \longrightarrow \mathcal{R}'$ als auch deren Umkehrung $f^{-1} : \mathcal{R}' \longrightarrow \mathcal{R}$ Bisimulationen sind. Nun erfüllt $\mathcal{R}$ im Gegensatz zu $\mathcal{R}'$ die Bedingung der Irreflexivität. Wenn wir daraus auf die Undefinierbarkeit der Irreflexivität schließen würden, dann wäre die Konklusion wahr. So weit, so gut. In umgekehrter Bisimulationsrichtung erfüllt $\mathcal{R}'$ im Gegensatz zu $\mathcal{R}$ die Bedingung der Reflexivität. Es ist klar, daß wir daraus nicht auf die Undefinierbarkeit der Reflexivität schließen dürfen, denn diese Bedingung *ist* definierbar – jedenfalls im bisher verwendeten Sinne von Definierbarkeit auf der Ebene von Rahmen:

> Eine Bedingung ist *(rahmen-)definierbar* gdw es eine Formel(menge) gibt, die in genau den Rahmen gültig ist, welche die Bedingung erfüllen.

Diagramme wie (a) weisen auf eine andere Art von Undefinierbarkeit hin. Das Diagramm zeigt, daß es keine Formel geben kann, welche an genau den Punkten eines Modells wahr ist, die reflexiv (bzw. irreflexiv) sind. Mit anderen Worten, es gibt keine Formel, die *punktweise* mit dieser Eigenschaft korrespondiert.

> Eine Bedingung ist *punktweise definierbar* gdw es eine Formel(menge) gibt, die an genau den Punkten eines Modells wahr ist, welche die Bedingung erfüllen.

Schon oben (p. 118) haben wir darauf hingewiesen, daß Reflexivität im punktweisen Sinne nicht definierbar sein kann. Die punktweise Korrespondenz zwischen einer Formel und einer Bedingung impliziert deren Korrespondenz auf der Ebene der Rahmen – aber nicht umgekehrt! Deshalb folgen aus Bisimulationsbeobachtungen wie in Diagramm (a) nicht unmittelbar Undefinierbarkeitsresultate auf der Rahmenebene. Die Bisimulationsinvarianz einer Eigenschaft ist keine notwendige Bedingung für ihre Definierbarkeit auf der Ebene der Rahmen; sie ist es jedoch auf der Ebene der Modelle, wie das folgende Korollar aus dem Bisimulationssatz zeigt.

KOROLLAR 44. *Wenn eine Klasse von Modellen modal definierbar ist, dann ist sie unter surjektiver Bisimulation abgeschlossen.*

BEWEIS. Wir nehmen an, daß eine Klasse $\mathbb{M}$ von Modellen durch eine Formelmenge X definiert ist, d.h. für alle Modelle $\mathcal{M}$,

$$(*) \qquad \mathcal{M} \in \mathbb{M} \text{ gdw } \mathcal{M} \models X.$$

Für *reductio* nehmen wir ferner an, daß $\mathbb{M}$ nicht unter surjektiver Bisimulation abgeschlossen ist.[23] Dann gibt es $\mathcal{M} \in \mathbb{M}$ und (†) $\mathcal{M}' \notin \mathbb{M}$ mit einer Bisimulation $H : \mathcal{M} \longrightarrow \mathcal{M}'$. Nach dem Bisimulationssatz 42 haben wir für alle Punkte a in $\mathcal{M}$ und a' in $\mathcal{M}'$: Wenn Haa', dann

$$a \models X \text{ in } \mathcal{M} \text{ gdw } a' \models X \text{ in } \mathcal{M}'.$$

Da H surjektiv ist und $\mathcal{M} \models X$, so folgt $\mathcal{M}' \models X$. Aber dann ist nach $(*)$ $\mathcal{M}' \in \mathbb{M}$ – entgegen unserer Annahme (†). ∎

Bisimulationen sind ein wichtiges Werkzeug um aus gegebenen Modellen modal äquivalente Modelle zu erzeugen. So kann ein Modell auf ein kleineres äquivalentes reduziert werden (*model contraction*), oder zu einem größeren mit gewünschten Struktureigenschaften erweitert werden (*tree unraveling*). Da wir uns mit solchen Verfahren im weiteren nicht beschäftigen werden, soll der bloße Hinweis auf solche Anwendungen genügen.

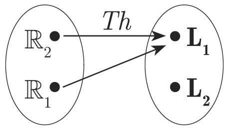
Undeterminierbare Logiken. Im vorigen Abschnitt haben wir festgestellt, daß es nicht für jede Rahmenklasse (bzw. Bedingung, die eine solche Klasse festlegt) eine diese definierende Logik (bzw. ein korrespondierendes Schema) gibt. Das folgt aus der Schlüsselbeobachtung, daß die Funktion Th nicht injektiv ist. Daraus folgt übrigens auch unmittelbar, daß es keine Umkehrfunktion für Th geben kann. Insbesondere gilt nicht

[23] Eine Bisimulation $H : W \longrightarrow W'$ ist *surjektiv*, wenn sie jedes Element in W' als Bild eines Arguments in W darstellt.

$Rm(Th(\mathbb{R})) = \mathbb{R}$ – die Gleichung $Th(Rm(\mathbf{L})) = \mathbf{L}$ gilt dagegen aufgrund der Definitionen. Die nächste naheliegende Frage ist:

· Gibt es für jede Logik eine sie determinierende Rahmenklasse? Bzw.:
· Gibt es für jedes Schema eine korrespondierende Rahmenbedingung?

Hier fragen wir danach, ob Th surjektiv ist, d.h. den Bereich normaler Modallogiken ausschöpft. Die Hypothese, daß das so sei, würde dieses Bild (mit undeterminierter Logik $\mathbf{L}_2$) ausschließen:



Die Hypothese ist falsch; das Bild ist richtig. Es gibt normale Modallogiken, im Bild $\mathbf{L}_2$, die von keiner Rahmenklasse determiniert werden – oder, anders gesagt, $Rm(\mathbf{L}_2) = \emptyset$.

Seit den 70er Jahren sind Logiken bekannt, welche auf $\mathbf{L}_2$ im Diagramm passen. Das erste solche System wurde von Thomason [280] in der Zeitlogik (mit den zwei Notwendigkeitsoperatoren **H** und **G**) gefunden; vgl. die Darstellung in [113, Kap. 7] und [31, Kap. 4.4]. Der Aufsatz [280] von Thomason enthält auch schon ein unvollständiges System in der Sprache $\mathcal{L}^{\Box}$. Die hier vorgestellte unvollständige Logik geht auf Boolos und Sambin (1985) zurück; vgl. [33, Kap. 11].

Es sei **KH** das System **K** +

H. $$\Box(\Box A \leftrightarrow A) \rightarrow \Box A,$$

und **KW** sei **K** +

W. $$\Box(\Box A \rightarrow A) \rightarrow \Box A.$$

(H steht für "Henkin", W für "(umgekehrt) wohlfundiert" – so jedenfalls wird das Schema W von Segerberg (1971) in [260] eingeführt. Aus einem Grund, den wir im nächsten Abschnitt kennenlernen werden, wird W auch "L", für "Löb", genannt.) Nun können wir folgendes beobachten:

1. H und W sind in genau denselben Rahmen gültig: $Rm(\mathrm{H}) = Rm(\mathrm{W})$.[24]
2. Das Schema 4: $\Box A \rightarrow \Box\Box A$ ist in **KW**, jedoch nicht in **KH** beweisbar.[25]

---

[24] Boolos [33, pp. 150ff.] gibt einen direkten Beweis. Die Beweise in [141, pp. 161ff.] und [81, pp. 206ff.] sind einfacher, setzen aber die Vollständigkeit von **KW** bzgl. einer Rahmenklasse voraus, die wir im nächsten Abschnitt einführen werden.

[25] Die erste Behauptung wird in Beobachtung 55 (s. n. Abschn.) bewiesen. Der Beweis der zweiten Behauptung stammt von Magari; Boolos [33, p. 152] stellt eine von Cresswell gefundene Vereinfachung vor.

Satz 45. *Das System* **KH** *wird von keiner Rahmenklasse determiniert.*

Beweis. Angenommen, es gäbe eine Rahmenklasse $\mathbb{H}$ mit $\mathbf{KH} = Th(\mathbb{H})$. Nach 2 enthält $\mathbb{H}$ dann einen Rahmen $\mathcal{R} \in Rm(\mathrm{H})$ mit $\mathcal{R} \not\models 4$. Aber nach 1 ist $\mathcal{R}$ auch in $Rm(\mathrm{W})$ und also (nach 2) $\mathcal{R} \models 4$ – Widerspruch. ■

Nicht-elementare Rahmenklassen. Bisher haben wir Rahmenbedingungen betrachtet, die sich in der Prädikatenlogik erster Stufe ausdrücken lassen: Alle Quantoren in einer solchen Bedingung binden nur Variablen, die sich auf die Elemente des Bereichs beziehen. Aber wir können natürlich auch Rahmenklassen durch Bedingungen höherer Stufe festlegen. Wenn eine so beschriebene Rahmenklasse sich nicht auch durch eine Menge erststufiger Formeln adäquat beschreiben läßt, dann sprechen wir von einer *nicht-elementaren Rahmenklasse.*

Ein einfaches Beispiel ist die Bedingung (Endl), daß der Bereich W eines Rahmens *endlich* sei. Damit meinen wir, daß es eine natürliche Zahl gibt so, daß der Bereich bis zu dieser Zahl vollständig abgezählt werden kann. Mit anderen Worten: Es gibt eine natürliche Zahl n und eine Bijektion f von $\{x \in \mathbf{N} : x \leq n\}$ nach W. Hier quantifizieren wir ($\exists f$!) über Relationen in $\mathbf{N} \times W$ – ersichtlich eine Quantifikation zweiter Stufe.

Ein weiteres Beispiel zweiter Stufe ist die Bedingung (Uwf), daß es im Bereich keine unendlich aufsteigenden R-Ketten gebe – äquivalent: in jeder nichtleeren Teilmenge von W gebe es maximale Elemente (unter R). *Umgekehrt wohlfundiert* (uwf) wird diese Bedingung in [260] genannt:

(Uwf) $\forall M \subseteq W : M \neq \emptyset \Rightarrow \exists a \in M(\neg \exists b \in M : Rab)$.

Nun haben wir eingangs bemerkt, daß wir in den modalen Operatoren Quantoren über die Trägermengen von Rahmen sehen können. Die Standardübersetzung übersetzt modale Formeln in prädikatenlogische Formeln erster Stufe. Wenn nun eine Rahmenklasse nicht-elementar, d.h. wesentlich durch eine Bedingung höherer Stufe festgelegt ist, dann wäre es überraschend, wenn modallogische Formeln eine solche Rahmenklasse definieren könnten. Aber genau dies geschieht im Falle der Klasse von Rahmen, die (Uwf) erfüllen. Denn das nächste Lemma besagt, daß das Schema W mit der nicht-elementaren Bedingung (Uwf) korrespondiert. Das Schema ist also ein Beispiel für eine Formel, die in gewisser Weise etwas über Rahmen sagt, was in der Sprache erster Stufe nicht gesagt werden kann.

Lemma 46. *Das Schema W korrespondiert mit der Bedingung (Uwf):* $Rm(W) = Uwf(\mathbb{K})$.

Beweis. Wir müssen für beliebige Rahmen $\mathcal{R} = (W, R)$ zeigen, daß

$$\mathcal{R} \models \Box(\Box A \to A) \to \Box A \text{ gdw } Uwf(\mathcal{R}).$$

(LR): Wir nehmen an, ein Rahmen sei nicht uwf und konstruieren darauf ein Modell mit einem Punkt, der W falsifiziert. Unter der Annahme gibt es eine unendlich aufsteigende R-Kette von Punkten $aRa'Ra''R....$ Sei K die Menge der Punkte in dieser Kette. Für jedes Atom P sei $I(P,x)=1$ gdw $x \notin K$. Dann gilt (1) $\forall x \in W : x \models \Box P \rightarrow P$. Denn wenn $x \notin K$, dann $x \models P$ (Def. I). Und wenn $x \in K$, dann gibt es ein $x' \in K$ mit Rxx' und $x' \not\models P$; also $x \not\models \Box P$. In beiden Fällen folgt (1), woraus weiter folgt, daß $\forall x \in W : x \models \Box(\Box P \rightarrow P)$. Für jedes $x \in K$ gilt jedoch, daß $x \not\models \Box P$. Also falsifizieren alle Punkte in K das Schema W.

(RL): Wir nehmen an, W sei in einem uwf'en Rahmen nicht gültig. Dann gibt es auf diesem Rahmen ein Modell mit einem Punkt a, so daß (1) $a \models \Box(\Box A \rightarrow A)$ und (2) $a \not\models \Box A$. Aus (2) folgt, daß es einen weiteren Punkt a' gibt mit Raa' und $a' \not\models P$. Aus (1) folgt, daß $a' \models \Box A \rightarrow A$. Aus beiden Annahmen zusammen folgt so, daß ($2'$) $a' \not\models \Box A$. Nun folgt aus ($2'$) wiederum, daß es ein a'' gibt mit $Ra'a''$ und $a'' \not\models P$. Mit (1) folgt daraus ($2''$) $a'' \not\models \Box A$. Das Argument generiert eine unendlich aufsteigende Kette $aRa'Ra''R\ldots$ – im Widerspruch zur Annahme (Uwf). ∎

Das Schema W definiert nicht nur die Klasse der uwf'en Rahmen, sondern ist als *Löb-Schema* auch das charakteristische Axiom der Gödel-Löb Logik **GL** – womit wir beim Thema des nächsten Abschnitts sind.

11. Notwendigkeit als Beweisbarkeit

Die Entwicklung der Modallogik in der zweiten Hälfte des 20. Jahrhunderts war nicht unumstritten. Die Skepsis bezog sich zum einen sehr allgemein auf den angeblich unklaren Begriff der Notwendigkeit, dessen formale Eigenschaften beschrieben werden sollten; zum anderen richteten sich spezifische Einwände gegen philosophisch kontroverse Folgerungen, welche die modale Prädikatenlogik nahezulegen schien. Zugleich war jedoch kaum zu bestreiten, daß in und mit der Modallogik wichtige Resultate erzielt wurden. Deshalb gab es schon früh in der modernen Modallogik Versuche, oder zumindest Andeutungen, den etwas schillernden Begriff der Notwendigkeit in einer klar verständlichen und unkontroversen Weise zu interpretieren. So beansprucht C.I. Lewis, besonders in seinen frühen Schriften, daß der Begriff einer strikten (d.h. notwendigen) Implikation so etwas wie Ableitbarkeit in einem axiomatischen System wiedergeben soll. [26]

Brouwer, Heyting und Kolmogorov hatten zum Verständnis der intuitionistischen Aussagenlogik **IL** vorgeschlagen, daß deren Junktoren keine

[26] Ganz ähnliche Stellen findet man übrigens später in den Arbeiten von Anderson und Belnap [10] zur Relevanzlogik: Relevante Implikation soll ein echtes Ableitbarkeitsverhältnis ausdrücken.

Wahrheits- sondern Beweisbarkeitsfunktionen seien: Eine Formel wie $A \vee B$ sei im Sinne von “A ist beweisbar oder B ist beweisbar” zu lesen (siehe z.B. [133]). In dem kurzen Aufsatz [112] aus dem Jahre 1933 geht Gödel – einer Anregung Parrys folgend – der Frage nach, welche modale Logik diese Interpretation von **IL** richtig wiedergibt, wenn wir den Operator $\Box$ im Sinne von “ist beweisbar” verstehen. Im Abschnitt 14 werden wir sehen, daß die Übersetzung von **IL** unter dieser Interpretation das modale System **S4** erzeugt. Jedoch schließt Gödel seine Notiz mit der folgenden Beobachtung:

> Es ist zu bemerken, daß für den Begriff “beweisbar in einem bestimmten formalen System S” die aus [**S4**] beweisbaren Formeln nicht alle gelten. Es gilt z.B. für ihn $\Box(\Box A \to A)$ niemals, d.h. für kein System S, das die Arithmetik enthält. Denn anderenfalls wäre beispielsweise $\Box(0 \neq 0) \to 0 \neq 0$ und daher auch $\neg\Box(0 \neq 0)$ in S beweisbar, d.h. die Widerspruchsfreiheit von S wäre in S beweisbar.

Gödel weist hier auf folgendes hin. Angenommen, wir verstehen die intuitionistischen Junktoren so, daß darin implizit von Beweisbarkeit die Rede ist. Wenn wir nun durch Übersetzung in eine modale Sprache diese Rede im Sinne von **IL** explizit machen, dann entsteht das System **S4**. Aber $\Box$ in **S4** hat Eigenschaften, die auf das arithmetisierte Beweisbarkeitsprädikat *Bew* im System **PA** der formalisierten Arithmetik nicht zutreffen. Beispielsweise kann $Bew(Bew A \to A)$ kein Theorem von **PA** sein. Anderenfalls könnten wir in **PA** einen Satz ableiten, der die Widerspruchsfreiheit von **PA** ausdrückt. Das können wir aber nach Gödels zweitem Unvollständigkeitssatz nicht. Also kann **S4** nicht die logischen Eigenschaften von *beweisbar-in-***PA** wiedergeben. Die Grundfrage für eine Beweisbarkeitsinterpretation der Modallogik können wir nun so formulieren: Welches modallogische System ist so, daß darin $\Box$ genau auf das Prädikat *Bew* in **PA** paßt?

Bevor wir diese Frage beantworten, sollten wir kurz erklären, warum die Theorie **PA** in diesem Zusammenhang eine so ausgezeichnete Rolle spielt und was das arithmetisierte Beweisbarkeitsprädikat *Bew* für **PA** ist. Wir müssen uns hier mit bloßen Hinweisen begnügen. Genauere Auskunft geben Darstellungen von Gödels Unvollständigkeitssätzen; siehe insbesondere die sehr zugänglichen Bücher von Smith [269, 270] oder auch [81].

Die sogenannte *Peano-Arithmetik*[27] **PA** ist eine Erweiterung der Prädikatenlogik erster Stufe (mit Identität) um arithmetische Axiome. Diese beschreiben ein Anfangsobjekt (0) sowie drei Operationen auf Zahlen: Nach-

[27] Eigentlich besser “Dedekind-Arithmetik”, wie wir heute wissen. Denn die sogenannten Peano-Axiome wurden zuvor (1888) schon von Dedekind vorgestellt.

folger ($'$), Addition und Multiplikation. Ferner enthält **PA** alle Instanzen eines Induktionsschemas.

Die erste, soeben gestellte Frage – warum ist **PA** eine so wichtige Theorie? – hat eine offensichtliche Antwort: Es handelt sich um die Axiomatisierung der Theorie der natürlichen Zahlen, aus der sich die Theorie aller weiteren Zahlbereiche entwickeln läßt. Durch Axiomatisierung erhalten wir einen präzisen Begriff von Ableitbarkeit.

Diese Anwort kann ergänzt werden um einen Aspekt, der uns auch zur Beantwortung der zweiten Frage führen wird. **PA** ist eine sehr einfache Theorie in einer sehr einfachen Sprache. Es gibt nur eine Individuenkonstante, die Null. Es gibt nur drei arithmetische Operationen: die Bildung des Nachfolgers einer Zahl, sowie die Addition und die Multiplikation. Das Induktionsschema dient der Ableitung von allgemeinen Aussagen über Zahlen. (Ohne das Schema könnten wir zwar für jeden Wert von x feststellen, daß $x + 0 = 0$, jedoch nicht die allgemeine Aussage $\forall x(x + 0 = 0)$ ableiten.) Obgleich **PA** so einfach ist, ist die Theorie erstaunlich stark!

- **PA** kann jede primitiv rekursive (p.r.) Funktion oder Eigenschaft von Zahlen richtig wiedergeben.[28] Im Falle einer p.r. Eigenschaft F ist damit gemeint, daß es ein Prädikat ϕx in der Sprache von **PA** gibt so, daß für jede Zahl n gilt: $\vdash \phi\bar{n}$, falls Fn, und $\vdash \neg\phi\bar{n}$, anderenfalls. (Hier, wie im folgenden, bezeichne $\vdash$ Ableitbarkeit in **PA**, und $\bar{n}$ sei irgendeine Darstellung der Zahl n in der Sprache von **PA**; also z.B. “$0'$” für die Zahl 1.)
- Wir können durch geeignete Kodierung Formeln und auch Folgen von Formeln (von denen einige Beweise in **PA** sind) eindeutig Zahlen zuordnen. Wenn A eine Formel ist, dann sei $\ulcorner A \urcorner$ die Kodierung von A als Zahl; $\ulcorner A \urcorner$ ist die *Gödelzahl* der Formel A.
- Die Eigenschaft einer geschlossenen Formel (Satz), in **PA** beweisbar zu sein, ist p.r.. Also ist auch die Eigenschaft einer Zahl, die Gödelzahl eines (in **PA**) beweisbaren Satzes zu sein p.r. und kann in **PA** richtig wiedergegeben werden.
- Es gibt daher ein Prädikat *Bew* in der Sprache von **PA** so, daß für jeden Satz A gilt: $\vdash Bew(\ulcorner A \urcorner)$, falls $\vdash A$, und $\vdash \neg Bew(\ulcorner A \urcorner)$, falls $\nvdash A$. *Bew* ist das *arithmetisierte Beweisbarkeitsprädikat* für **PA**.

Stark ist **PA** also nicht nur im Hinblick auf die Menge ableitbarer arithmetischer Sätze, sondern auch darin, auf beliebige Ausdrücke und Folgen solcher

[28] “Primitiv rekursiv”: Es reicht für unsere Zwecke an Funktionen zu denken, die in einem intuitiven Sinne mechanisch berechenbar sind, bzw. an Eigenschaften, von denen in diesem Sinne entscheidbar sind, ob sie zutreffen oder nicht.

Ausdrücke Bezug zu nehmen und p.r. Eigenschaften solcher Ausdrücke richtig wiedergeben zu können.

Über *Bew* in **PA** wissen wir einiges. Zu den Schlüsseleigenschaften von *Bew* gehören die folgenden von Hilbert und Bernays [134] herausgestellen und von Löb [190] eleganter formulierten *Ableitbarkeitsbedingungen*:[29]

HBL1 $\vdash A \quad \Rightarrow \quad \vdash Bew(\ulcorner A \urcorner)$.
HBL2 $\vdash Bew(\ulcorner A \to B \urcorner) \to (Bew(\ulcorner A \urcorner) \to Bew(\ulcorner B \urcorner))$.
HBL3 $\vdash Bew(\ulcorner A \urcorner) \to Bew(\ulcorner Bew(\ulcorner A \urcorner) \urcorner)$.

HBL1 ist natürlich eine Hälfte der genannten Eigenschaft von *Bew*, zwischen ableitbaren und nicht ableitbaren Sätzen richtig zu unterscheiden. Da wir auch die andere Hälfte benötigen werden, fügen wir sie der HBL-Liste kurzerhand hinzu (obwohl sie gewöhnlich nicht zu den HBL-Bedingungen gezählt wird):

HBL0 $\nvdash A \quad \Rightarrow \quad \vdash \neg Bew(\ulcorner A \urcorner)$.

Eine weitere wichtige Eigenschaft ergibt sich aus dem

Diagonalisierungslemma 47. *Es sei ϕx eine (nur für x) offene Formel in der Sprache von* **PA**. *Dann gibt es einen Satz so, daß $\vdash D \leftrightarrow \phi(\ulcorner D \urcorner)$. ($D$ heißt ein* Fixpunkt *für ϕ.)*[30]

Wenn wir das Lemma auf $Bew(x)$ bzw. $\neg Bew(x)$ anwenden, dann erhalten wir *Henkin-* bzw. *Gödel-Fixpunkte*:

HFP $\exists H\colon \vdash H \leftrightarrow Bew(\ulcorner H \urcorner)$, und
GFP $\exists G\colon \vdash G \leftrightarrow \neg Bew(\ulcorner G \urcorner)$.

Unvollständigkeit, modallogisch. Nach diesen Sondierungen über Beweisbarkeit in **PA** kehren wir zur Modallogik zurück. Wenn wir $Bew(\ulcorner A \urcorner)$ als $\Box A$ schreiben, dann werden aus den Bedingungen HBL1–3 für *Bew* vertraute Prinzipien der Modallogik, nämlich:

RN. $$\frac{A}{\Box A}$$

K. $$\Box(A \to B) \to (\Box A \to \Box B)$$

4. $$\Box A \to \Box\Box A$$

[29] Den Nachweis der HBL-Bedingungen für **PA** werden wir hier nicht führen. Er ist zwar nicht besonders schwierig, trägt aber zum Verständnis des Themas in diesem Abschnitt wenig bei. Außer in dem gut lesbaren Aufsatz [190] von Löb, findet sich der Beweis in beinahe jeder Darstellung der Gödel-Sätze.

[30] Beweis siehe z.B. [270, Kap. 14] oder [34, Kap. 17].

Im modalen System **K4** erfüllt $\Box$ also die HBL-Ableitungsbedingungen 1–3 für *Bew*. Um die Beziehung zwischen den HBL-Bedingungen und den modalen Postulaten genauer zu fassen, wollen wir den Begriff einer *Interpretation* der modalen Sprache $\mathcal{L}^{\Box}$ von **K4** in der arithmetischen Sprache $\mathcal{L}^{\sharp}$ von **PA** definieren. Eine solche Interpretation * ordnet jedem Atom P in $\mathcal{L}^{\Box}$ einen arithmetischen Satz P^* in $\mathcal{L}^{\sharp}$ zu. Auf der Basis einer solchen Zuordnung werden alle weiteren Formeln so interpretiert:

$$\begin{aligned} \bot^* &= (0 = 1) \\ (\neg A)^* &= \neg A^* \\ (A \wedge B)^* &= A^* \wedge B^* \\ (\Box A)^* &= Bew(\ulcorner A^* \urcorner) \end{aligned}$$

Dann stehen **K4** und **PA** in der folgenden Beziehung zueinander:

Satz 48. (Arithmetische Richtigkeit von **K4**.) *Wenn $A \in$ **K4**, dann $A^* \in$ **PA** für alle Interpretationen *.*

Beweis. Es gilt die Axiome und Regeln von **K4** unter einer beliebig gewählten Interpretation in **PA** zu verifizieren. Sei * also eine solche Interpretation. Jede Tautologie ist auch unter * eine Tautologie und also in **PA**. Modus Ponens überträgt Beweisbarkeit auch in **PA**. Die modalen Postulate RN, K und 4 werden unter * in die HBL(1–3)-Bedingungen übersetzt, welche für *Bew* in **PA** gelten. ∎

Nach Satz 48 können wir nun modallogisch in **K4** argumentieren, um Einsichten über *Bew* in **PA** zu gewinnen. Wir illustrieren das in den Beweisen der folgenden Sätze. Der erste Satz ist eine modale Version der einen Hälfte von Gödels erstem Unvollständigkeitssatz: Aus der Annahme eines Gödel-Fixpunktes G folgt, daß falls $\bot$ nicht beweisbar ist, auch G nicht beweisbar ist. (Die andere Hälfte von Gödels erstem Satz besagt, daß unter der Annahme der ω-Konsistenz von **PA** auch $\neg G$ in **PA** nicht beweisbar ist.) In dieser Version soll also $\neg\Box\bot$ die Annahme der (einfachen, nicht der ω-) Widerspruchsfreiheit von **PA** in Gödels Satz ausdrücken.

Satz 49. $G \leftrightarrow \neg\Box G \vdash \neg\Box\bot \rightarrow \neg\Box G$ *in* **K4**.

Beweis. (Hier, wie in einigen der folgenden Beweise, verweist die Rechtfertigung "**K**" auf leicht nachvollziehbare Schritte, welche schon die modale Basislogik **K** zur Verfügung stellt.)

(1)	$\Box G \to \neg G$	aus der Annahme
(2)	$\Box\Box G \to \Box\neg G$	1, RN, K
(3)	$\Box G \to \Box\neg G$	2, Schema 4
(4)	$\Box G \to \Box G$	Taut.
(5)	$\Box G \to \Box G \wedge \Box\neg G$	3,4
(6)	$\Box G \to \Box(G \wedge \neg G)$	5, **K**
(7)	$\Box G \to \Box\bot$	6,$\bot$
(8)	$\neg\Box\bot \to \neg\Box G$	7

■

Aus diesem Satz dürfen wir jetzt so schließen: Da es in **PA** einen Gödel-Fixpunkt G gibt, so haben wir nach Satz 48 (kontraponiert),

$$\mathbf{PA} \vdash Bew(\ulcorner G \urcorner) \to Bew(\ulcorner 0 = 1 \urcorner). \tag{1}$$

Nun nehmen wir an, **PA** sei konsistent. Also

$$\mathbf{PA} \nvdash 0 = 1. \tag{2}$$

Nach HBL0 folgt aus (2), $\mathbf{PA} \vdash \neg Bew(\ulcorner 0 = 1 \urcorner)$ und desweiteren (Konsistenz!)

$$\mathbf{PA} \nvdash Bew(\ulcorner 0 = 1 \urcorner). \tag{3}$$

Aber dann aus (1) und (3)

$$\mathbf{PA} \nvdash Bew(\ulcorner G \urcorner).$$

Wenn nun $\mathbf{PA} \vdash G$, dann hätten wir auch $\mathbf{PA} \vdash Bew(\ulcorner G \urcorner)$ – Widerspruch. Also folgt unter der Annahme, daß **PA** konsistent ist, daß G in **PA** nicht ableitbar ist – die eine Hälfte von Gödels erstem Unvollständigkeitssatz.

Der zweite Unvollständigkeitssatz besagt, daß wenn **PA** konsistent ist, dann ist der Satz, daß **PA** konsistent ist – nämlich $\neg Bew(\ulcorner 0 = 1 \urcorner)$ – in **PA** nicht beweisbar. In der modalen Sprache können wir das so ausdrücken: Wenn $\neg\Box\bot$, dann $\neg\Box(\neg\Box\bot)$. Dann wird als Annahme nur noch die Existenz eines Gödel-Fixpunktes G benötigt.

Satz 50. *$G \leftrightarrow \neg\Box G \vdash \neg\Box\bot \to \neg\Box\neg\Box\bot$ in* **K4**.

Beweis.

(1)	$G \leftrightarrow \neg\Box G$	Annahme
(2)	$\neg\Box\bot \to \neg\Box G$	1, Satz 49
(3)	$\neg\Box G \to G$	1
(4)	$\neg\Box\bot \to G$	2,3
(5)	$\Box\neg\Box\bot \to \Box G$	4, RN, K
(6)	$\neg\Box G \to \neg\Box\neg\Box\bot$	5
(7)	$\neg\Box\bot \to \neg\Box\neg\Box\bot$	2,6

■

Um aus diesem Satz, den nichtmodalen zweiten Unvollständigkeitssatz zu gewinnen, argumentieren wir wie folgt. Da es in **PA** einen Gödel-Fixpunkt G gibt, können wir durch beliebige Übersetzung das Konsequens der Sequenz ablösen:

$$\mathbf{PA} \vdash \neg Bew(\ulcorner 0 = 1 \urcorner) \to \neg Bew(\ulcorner \neg Bew(\ulcorner 0 = 1 \urcorner) \urcorner).$$

Das Antezedens ist gegeben. Daher

$$\mathbf{PA} \vdash \neg Bew(\ulcorner \neg Bew(\ulcorner 0 = 1 \urcorner) \urcorner).$$

Angenommen – für *reductio* –, $\mathbf{PA} \vdash \neg Bew(\ulcorner 0 = 1 \urcorner)$. Nach HBL1 folgt daraus sofort $\mathbf{PA} \vdash Bew(\ulcorner \neg Bew(\ulcorner 0 = 1 \urcorner) \urcorner)$ – im Widerspruch zur angenommenen Konsistenz von **PA**. Unter der Annahme gilt also $\mathbf{PA} \nvdash \neg Bew(\ulcorner 0 = 1 \urcorner)$.

Schließlich wollen wir noch zeigen, daß alle Gödel-Fixpunkte in **PA** äquivalent sind zur Konsistenzaussage $\neg Bew(\ulcorner 0 = 1 \urcorner)$ – modal gewendet:

Satz 51. $G \leftrightarrow \neg \Box G \vdash G \leftrightarrow \neg \Box \bot$ *in* **K4**.

Beweis. $G \leftrightarrow \neg \Box G \vdash \neg \Box \bot \to G$ wurde in Zeile (4) des Beweises von Satz 50 gezeigt. Die andere Implikationsrichtung folgt so:

(1)	$\bot \to G$	Aussagenlogik
(2)	$\Box(\bot \to G)$	1, RN
(3)	$\Box\bot \to \Box G$	2, K
(4)	$\neg\Box G \to \neg\Box\bot$	3
(5)	$G \to \neg\Box\bot$	4, $\neg\Box G \leftrightarrow G$

■

Der Satz erlaubt einen kurzen Weg vom ersten zum zweiten Unvollständigkeitssatz: Da Gödel-Fixpunkte G in **PA** (unter der Annahme der Widerspruchsfreiheit) nicht ableitbar sind (erster Satz), so kann auch die Konsistenzaussage $\neg Bew(\ulcorner 0 = 1 \urcorner)$ nicht ableitbar sein (zweiter Satz). Und umgekehrt können wir von $\mathbf{PA} \nvdash \neg Bew(\ulcorner 0 = 1 \urcorner)$ auf $\mathbf{PA} \nvdash G$ schließen, falls **PA** konsistent ist.

Löbs Satz und GL. Im letzten Abschnitt fiel Gödel-Fixpunkten für das Prädikate $\neg Bew$ eine Schlüsselrolle zu. Solche Fixpunkte sind in **PA** nicht ableitbar. Aber wie steht es mit Henkin-Fixpunkten H für das Prädikat Bew? Sind diese ableitbar? Die Frage stellte Leon Henkin und die Antwort folgt aus

LÖBS SATZ 52. (Löb [190].) *Wenn* **PA** $\vdash$ $Bew(\ulcorner A \urcorner) \to A$, *dann* **PA** $\vdash$ A.

Setzen wir H für A ein, dann ist das Antezedens durch die Fixpunktäquivalenz HFP gegeben; also ist H in **PA** beweisbar. Wir können Löbs Satz aus der folgenden Beobachtung über **K4** beweisen:

LEMMA 53. $L \leftrightarrow (\Box L \to A), \Box A \to A \vdash A$ *in* **K4**.

BEWEIS. (Der Beweis aus der Annahme $L \leftrightarrow (\Box L \to A)$ und mit der Anwendung von Kontraktion in Zeile 7 ist der Argumentation in der Paradoxie von Curry [55] sehr ähnlich.)

(1)	$L \leftrightarrow (\Box L \to A)$	Annahme
(2)	$\Box L \to \Box(\Box L \to A)$	1, RN, K
(3)	$\Box L \to (\Box\Box L \to \Box A)$	2, K
(4)	$\Box\Box L \to (\Box L \to \Box A)$	3
(5)	$\Box L \to \Box\Box L$	Schema 4
(6)	$\Box L \to (\Box L \to \Box A)$	4,5
(7)	$\Box L \to \Box A$	6
(8)	$\Box A \to A$	Annahme
(9)	$\Box L \to A$	7,8
(10)	$(\Box L \to A) \to A$	1
(11)	A	9,10

■

Löbs Satz erhalten wir nun so: Nach dem Diagonalisierungslemma hat jede Formel $Bew(x) \to A$ einen Fixpunkt in **PA**, d.h. für jeden Satz A gibt es einen Satz L so, daß **PA** $\vdash L \leftrightarrow (Bew(\ulcorner L \urcorner) \to A)$. **PA** erfüllt somit die erste Annahme im Lemma und Löbs Satz folgt nach Übersetzung.

Gödels zweiter Unvollständigkeitssatz folgt als ein Korollar zu Löbs Satz. Man setze $\bot$ für A ein. Dann folgt nach Kontraposition und Umformung von $Bew(\ulcorner \bot \urcorner) \to \bot$ zu $\neg Bew(\ulcorner \bot \urcorner)$:

$$\text{Wenn } \mathbf{PA} \nvdash \bot, \text{ dann } \mathbf{PA} \nvdash \neg Bew(\ulcorner \bot \urcorner).$$

Wie die HBL-Bedingungen 1–3, so charakterisiert auch Löbs Satz das Beweisbarkeitsprädikat in **PA**:

HBL4 $\quad \vdash Bew(\ulcorner A \urcorner) \to A \quad \Rightarrow \quad \vdash A.$

Diese Bedingung entspricht der modalen Regel

LR. $$\frac{\Box A \to A}{A}.$$

K4 erlaubt uns nicht im allgemeinen nach dieser Regel zu schließen. Lemma 53 führt den Schluß unter einer Bedingung vor, nämlich der weiteren Prämisse, daß es einen Löb-Fixpunkt gibt, d.h. der Äquivalenz $L \leftrightarrow (\Box L \to A)$. In **PA** folgt diese aus dem Diagonalisierungslemma. Wenn wir **K4** um die Regel LR erweitern, dann nähern wir uns offenbar weiter dem Ziel *Bew* für Theorien wie **PA** zu charakterisieren, für die ein solches Diagonalisierungslemma zur Verfügung steht.

Anmerkung. Tatsächlich hat die Hinzunahme von W genau den Effekt, daß nun beliebige, in der modalen Sprache ausdrückbare Eigenschaften diagonalisiert werden können. Das ist der Inhalt des Fixpunktsatzes von de Jongh und Sambin:[31]

FIXPUNKTSATZ 54. *Es sei $A[P]$ eine Formel, in der jedes Vorkommen von P in A im Skopus eines Vorkommens von $\Box$ steht. Dann gibt es eine Formel F so, daß $\vdash F \leftrightarrow A[F/P]$ in* **K4**+LR, *wobei P in F nicht vorkommt, und in F nur $\top$ oder solche Atome vorkommen, die auch in A vorkommen. (F ist ein Fixpunkt für $A[P]$.)*

Aus dem Beweis des Satzes ergibt sich ein Verfahren für die Lösung von Fixpunktäquivalenzen. *Beispiele*:

$A = \Box P$. Lösung $F = \top$: $\vdash \top \leftrightarrow \Box\top$ (Henkin-Fixpunkt!).

$A = \neg\Box P$. Lösung $F = \neg\Box\bot$: $\vdash \neg\Box\bot \leftrightarrow \neg\Box\neg\Box\bot$ (Gödel-Fixpunkt!).

$A = \Box P \to Q$. Lösung $F = \Box Q \to Q$: $\vdash (\Box Q \to Q) \leftrightarrow (\Box(\Box Q \to Q) \to Q)$ (Löb-Fixpunkt!).

Als Entsprechung zum Satz 51 erweisen sich alle Lösungen für eine gegebene Fixpunktäquivalenz als äquivalent in **K4**+LR. (*Ende der Anmerkung.*)

Wir können **K4**+LR auch auf andere Weise darstellen, nämlich als die Erweiterung der Basislogik **K** um das Löb-Schema

W. $$\Box(\Box A \to A) \to \Box A.$$

In der Bezeichnungsweise nach Chellas [53] erhalten wir so das System **KW**. In der Literatur über Beweisbarkeitslogik wird dieses System üblicherweise **GL** (für Gödel und Löb) genannt.

[31] Den Satz haben de Jongh und Sambin unabhängig voneinander gefunden. Für einen Beweis siehe [33, Kap. 8], [242], [81, pp. 199ff], oder auch [271, pp. 109ff] mit interessanten Hinweisen.

Lemma 55. *Die Mengen der Theoreme von* **K4**+LR *und* **GL** *sind gleich; insbesondere ist* $\Box A \to \Box\Box A \in \mathbf{GL}$.

Beweis. Die Richtung $\mathbf{GL} \subseteq \mathbf{K4} + \mathrm{LR}$ ist einfach. Für die umgekehrte Richtung ist die Verifikation des Schemas 4 in **GL** interessant. Um die Formeln im Beweis übersichtlicher zu machen, sei $\triangle A = \Box A \wedge A$.

(1)	$A \to (\Box\Box A \wedge \Box A \to \Box A \wedge A)$	Aussagenlogik
(2)	$A \to (\Box(\Box A \wedge A) \to \Box A \wedge A)$	1, **K**
(3)	$A \to (\Box\triangle A \to \triangle A)$	2, $\triangle$
(4)	$\Box A \to \Box(\Box\triangle A \to \triangle A)$	3, **K**
(5)	$\Box(\Box\triangle A \to \triangle A) \to \Box\triangle A$	W ($\triangle A$ für A)
(6)	$\Box A \to \Box\triangle A$	4,5
(7)	$\Box A \to \Box(\Box A \wedge A)$	6, $\triangle$
(8)	$\Box A \to \Box\Box A$	7, **K**

∎

Es bleibt noch zu zeigen, daß **GL** eine *echte* Erweiterung von **K4** ist. Wir wissen, daß **K4** die Klasse der transitiven Rahmen definiert. Es wird also genügen, das Löb-Schema $\Box(\Box A \to A) \to \Box A$ in einem Modell auf einem solchen Rahmen zu falsifizieren. Dazu betrachten wir ein Modell $(\{a, b\}, R, I)$ mit Rab und Rbb. Die Relation ist transitiv. Sei I so, daß $I(P, b) = 0$, im Bild:

Da $a \not\models \Box P$ (denn Rab und $b \not\models P$), so haben wir $a \models \Box P \to P$. Das gleiche gilt für b: $b \not\models \Box P$ (denn Rbb und $b \not\models P$) und also $b \models \Box P \to P$. Also $a \models \Box(\Box P \to P)$.

Satz 56. (Arithmetische Richtigkeit von **GL**). *Wenn* $A \in \mathbf{GL}$, *dann* $A^* \in \mathbf{PA}$ *für alle Interpretationen* *.

Beweis. Angesichts der Beobachtung 55 genügt es, den Beweis von Satz 48 um die Verifikation der Regel LR in **PA** zu erweitern. Angenommen $\mathbf{PA} \vdash (\Box A \to A)^*$, d.h. $\mathbf{PA} \vdash \mathit{Bew}(A^*) \to A^*$. Dann folgt nach Löbs Satz $\mathbf{PA} \vdash A^*$ wie gewünscht. ∎

Es gilt auch die Umkehrung:

Satz 57. (Arithmetische Vollständigkeit von **GL**, Solovay 1976). *Wenn* $A \notin \mathbf{GL}$, *dann gibt es eine Interpretation* * *so, daß* $A^* \notin \mathbf{PA}$.

Der Beweis würde den Rahmen unserer Darstellung sprengen; siehe [33, Kap. 9].

GL und Gödel-Rahmen. Es sei in Erinnerung gerufen, daß das System **GL** (= **KW**) durch folgende Schemata festgelegt ist:

τ.	Jede Tautologie
K.	$\Box(A \to B) \to (\Box A \to \Box B)$
W.	$\Box(\Box A \to A) \to \Box A$
RN.	$\dfrac{A}{\Box A}$
MP.	$\dfrac{A \quad A \to B}{B}$

Nach dem Korrespondenzlemma 46 könnte man versucht sein, auf die Richtigkeit und Vollständigkeit von **GL** bezüglich der Klasse der umgekehrt wohlfundierten (uwf'n) Rahmen zu schließen. Aber dieser Schluß setzt voraus, daß W ein *kanonisches* Schema ist; siehe dazu die ausführliche Diskussion auf pp. 118ff. Die Voraussetzung trifft nicht zu: Der Umstand, daß eine normale Modallogik W als Theorem enthält, garantiert nicht, daß der kanonische Rahmen für diese Logik uwf ist – im Falle von **GL** ist er es nicht.

Im Vollständigkeitsargument wollen wir für jedes Nicht-Theorem F von **GL** ein falsifizierendes Modell aufweisen, das zumindest uwf ist. Nach dem gerade Gesagten, kann das kein Modell auf dem kanonischen Rahmen für **GL** sein. Wir müssen solche Modelle daher etwas anders konstruieren. Es lohnt sich auch unabhängig vom Interesse an **GL**, die Art der Konstruktion zu kennen, denn sie kann ebenso auf andere normale Modallogiken angewandt werden, um deren Richtigkeit und Vollständigkeit bezgl. *endlicher* Rahmenklassen nachzuweisen.[32]

Wenn A eine Formel ist, dann sei $\mathrm{ntf}(A)$ die Menge, der in A vorkommenden Teilformeln zusammen mit deren Negationen.

- Eine Menge $X \subseteq \mathrm{ntf}(A)$ ist *konsistent*, wenn $X \nvdash \bot$ (in **GL**); und sie ist eine *A-Welt*, wenn X konsistent ist und für alle $B \in \mathrm{ntf}(A)$ gilt: entweder $B \in X$ oder $\neg B \in X$.[33]

Offensichtlich kann jede konsistente Teilmenge einer Menge $\mathrm{ntf}(A)$ zu einer A-Welt vervollständigt werden. Für eine gegebene Formel A sei W_A die

[32] In [34, 44, 81] wird das vorgeführt.

[33] Offensichtlich ist nicht jede A-Welt – im jetzt definierten Sinne – eine Menge, die A enthält. Andere Autoren nennen A-Welten maximal konsistente oder vervollständigte Teilmengen von $\mathrm{ntf}(A)$.

Menge der A-Welten. Für jede Menge $X \in W_A$ und jede Formel $B \in \mathrm{ntf}(A)$ gilt:

$$(*) \qquad X \vdash B \text{ gdw } B \in X.$$

(Die RL-Richtung ist trivial. Für die andere Richtung nehme man an, daß $X \vdash B$. Da $B \in \mathrm{ntf}(A)$, ist entweder $\neg B \in X$ oder $B \in X$. Der erste Fall ist ausgeschlossen, denn dann wäre X inkonsistent. Also muß $B \in X$ sein.)

Zunächst werden wir die Vollständigkeit von GL für Gödel-Rahmen beweisen.

- Ein *Gödel-Rahmen* (W, R) ist ein Kripke-Rahmen mit endlicher (nichtleerer) Menge W, in der R irreflexiv und transitiv ist. Die Klasse der Gödel-Rahmen notieren wir mit $\mathit{Gdl}(\mathbb{K})$.

Danach zeigen wir, daß die Gödel-Rahmen genau die endlichen, transitiven und uwf'en Rahmen sind. Als Korollar folgt ein zweiter Vollständigkeitssatz.

SATZ 58. (Vollständigkeit I, Segerberg [260].) *Alle in Gödel-Rahmen gültigen Formeln sind Theoreme von* **GL**: $\mathit{Th}(\mathit{Gdl}(\mathbb{K})) \subseteq \mathbf{GL}$.

BEWEIS. Es sei E eine in **GL** nicht ableitbare Formel. Dann können wir $\neg E$ konsistent zu **GL** hinzufügen. Wir konstruieren ein Modell auf einem Rahmen in $\mathit{Gdl}(\mathbb{K})$, welches einen Punkt a enthält so, daß $a \not\models E$. Im folgenden sei $\neg E$ mit F bezeichnet.

Im Modell (W_F, R_F, I_F) sei W_F die Menge der F-Welten wie oben angegeben. (Wir verwenden $a, x, y, z, ...$ für Elemente in W_F) Da $\mathrm{ntf}(F)$ endlich ist, so ist auch W_F *endlich.* Da $\{F\}$ konsistent ist, so kann $\{F\}$ zu einer F-Welt vervollständigt werden. Also ist $W_F \neq \emptyset$. Sei a eine solche F-Welt. Dann ist $\neg F \notin a$, d.h. a falsifiziert E. Wir setzen

$$I_F(P, x) = 1 \text{ gdw } P \in X.$$

Die Relation R_F (ab hier einfach R) sei so definiert:

$$Rxy \text{ gdw } \begin{cases} \text{(a) } \forall \Box A \in x\colon A, \Box A \in y, \text{ und} \\ \text{(b) } \exists \Box A \in y\colon \Box A \notin x. \end{cases}$$

Bedingung (b) erzwingt die *Irreflexivität* von R. Die *Transitivität* der Relation ist ebenfalls einfach nachzuprüfen.

Es bleibt zu zeigen, daß

$$\Box A \in x \quad \text{gdw} \quad \forall y : Rxy \Rightarrow y \models A.$$

Die LR-Richtung ist trivial. Für die andere Richtung nehmen wir an, daß

(1) $$\Box A \notin x.$$

Wir müssen nun die Negation der rechten Seite zeigen. Wenn wir diese nach der Definition von R auspacken, so erhalten wir: $\exists y \in W_F$ so, daß

(i) $\neg A \in y$,
(ii) $(\forall B)\ \Box B \in x \ \Rightarrow\ B, \Box B \in y$, und
(iii) $(\exists C)\ \Box C \in y$ und $\Box C \notin x$.

Für C in (iii) können wir A aus (1) einsetzen und erhalten so:

(iii′) $\Box A \in y$.

Um die Behauptung $\exists y \in W_F$...(i-iii′) zu beweisen, genügt es zu zeigen, daß diese Menge konsistent ist:

$$\{\neg A, \Box A\} \cup \{B, \Box B : \Box B \in x\}.$$

Angenommen, das sei nicht so, d.h.

(2) $$\neg A, \Box A, \{B, \Box B : \Box B \in x\} \vdash \bot.$$

Dann ...

(3)	$\{B \wedge \Box B : \Box B \in x\} \vdash \Box A \to A$	2, Aussagenlogik
(4)	$\{\Box(B \wedge \Box B) : \Box B \in x\} \vdash \Box(\Box A \to A)$	3, **K**
(5)	$\{\Box(B \wedge \Box B) : \Box B \in x\} \vdash \Box A$	4, W
(6)	$\{\Box B \wedge \Box\Box B : \Box B \in x\} \vdash \Box A$	5, **K**
(7)	$\Box B \to \Box\Box B$	**GL** (Lemma 55)
(8)	$\{\Box B : \Box B \in x\} \vdash \Box A$	6,7, **K**
(9)	$\{\Box B : \Box B \in x\} \subseteq x$	
(10)	$x \vdash \Box A$	8,9
(11)	$\Box A \in x$ ...	10, $(*)$

... im Widerspruch zu (1). ■

Lemma 59. *1. Jeder uwf'e Rahmen ist irreflexiv. 2. Jeder irreflexive Rahmen, der endlich und transitiv ist, ist uwf.*

Beweis. *Ad* 1. Es sei R uwf. Sei $x \in W$ beliebig gewählt. Dann ist x in $\{x\}$ maximal. Also $\forall x \in W\colon \neg\exists y(Rxy)$.

Ad 2. Es sei R nicht uwf. Dann gibt es eine nichtleere Teilmenge $M \subseteq W$, die kein maximales Element enthält. M bildet also eine unendlich aufsteigende Kette $x_0 R x_1 R x_2 R \ldots$. Aus den Annahmen der Transitivität und Irreflexivität folgt nun, daß alle Elemente in M distinkt sein müssen: Wenn R transitiv ist, gilt $Rx_i x_j$ für alle i, k mit $0 \leq i < k$. Und wenn R irreflexiv ist, gilt $x_i \neq x_j$ für jedes Paar (i, j). Also ist $M \subseteq W$ – im Widerspruch zu unserer Annahme – unendlich. ■

Aus dem Lemma folgt unmittelbar, daß die Gödel-Rahmen mit den endlichen, transitiven und uwf'en Rahmen koinzidieren. So folgt weiter aus Satz 58:

SATZ 60. (Vollständigkeit II.) *Alle in endlichen, transitiven und uwf'en Rahmen gültigen Formeln sind Theoreme von* **GL**.

KOROLLAR 61. *Die folgenden Aussagen sind äquivalent:*
1. A ist ein Theorem von **GL**.
2. A ist gültig in allen endlichen, transitiven und uwf'en Rahmen.
3. A ist gültig in allen Gödel-Rahmen.

BEWEIS. 1 $\Rightarrow$ 2: Nach Lemma 46 ist $\mathbf{GL} \subseteq Th(Uwf(\mathbb{K}))$. Also sind die Theoreme von GL auch in uwf'en Rahmen unter weiteren Bedingungen gültig. 2 $\Rightarrow$ 3: Lemma 59. 3 $\Rightarrow$ 1: Satz 58. ∎

Diese Resultate können verschärft werden zu Vollständigkeit in engeren Rahmenklassen. Aber es ist an der Zeit, zu weiteren Themen in der Modallogik überzugehen.

12. Ein einfaches Verfahren: Tableaus

Gegeben eine Logik **L** und eine Formel A, so ist die naheliegendste Frage überhaupt: Ist A ein Theorem von **L**?

Wenn wir über **L** nicht mehr wissen, als daß diese Logik durch bestimmte Axiome und Regeln definiert ist, dann bleibt uns nichts anderes übrig, als daran zu gehen, einen Beweis von A in **L** vorzulegen. Das kann ein sehr mühseliges Unterfangen sein. Wenn wir dagegen wissen, daß die Theoreme von **L** genau die Formeln sind, die in einer bestimmten Rahmenklasse gültig sind, dann ist eine Formel genau dann ein Theorem, wenn jeder Versuch, ein Gegenmodell auf Rahmen dieser Klasse zu konstruieren, scheitert. Diese Überlegung liegt schon den einfachen Verfahren zugrunde, die wir auch in der nicht-modalen Aussagenlogik verwenden, um zu prüfen, ob eine Formel ein Theorem (Tautologie) ist – z.B. Tableaus und deren Kurzform, das Quine'sche Verfahren.

Hier wollen wir die bekannten Tableau-Regeln für die klassische Aussagenlogik auf recht einfache Weise erweitern. Tableau-Regeln zielen darauf, möglichst rasch und systematisch ein Gegenmodell zu einer gegebenen Formel (oder Formelmenge) zu finden. Läßt sich kein Gegenmodell finden, dann ist die Negation der zu prüfenden Formel gültig in der Rahmenklasse, die dem Verfahren jeweils zugrundeliegt. Gegeben ein entsprechendes Vollständigkeitsresultat, können wir daraus auf die Beweisbarkeit der

Formel schließen. Der wesentliche Unterschied zu Tableaus für die nichtmodale Aussagenlogik besteht darin, daß wir nun die Knoten mit Punkten aus einer Modellmenge ("Welten") indizieren. Die Modaloperatoren wirken wie Quantoren über diese Indizes – übrigens, völlig analog zur Behandlung der Quantoren in Tableau-Verfahren für die Prädikatenlogik, wo Variablenbelegungen die Rolle der Indizes spielen.

Jede Anwendung des Tableau-Verfahrens bezieht sich immer auf einen gegebenen Rahmen (W, R). Für jede Formel A drückt eine Knotenbeschriftung $a : A$ aus, daß A am Punkt $a \in W$ wahr ist. Boole'sche Knoten erhalten Nachfolger nach diesen Regeln:

Tbl (Boole)

$$\begin{array}{c} a : A \wedge B \\ | \\ a : A \\ a : B \end{array} \qquad \begin{array}{c} a : \neg(A \wedge B) \\ / \qquad \backslash \\ a : \neg A \quad a : \neg B \end{array} \qquad \begin{array}{c} a : \neg\neg A \\ | \\ a : A \end{array}$$

Regeln für weitere Junktoren ergeben sich aus deren jeweiliger Definition. Für $\vee$ und $\rightarrow$ erhalten wir so diese Regeln:

$$\begin{array}{c} a : A \vee B \\ / \qquad \backslash \\ a : A \quad a : B \end{array} \qquad \begin{array}{c} a : \neg(A \vee B) \\ | \\ a : \neg A \\ a : \neg B \end{array} \qquad \begin{array}{c} a : A \rightarrow B \\ / \qquad \backslash \\ a : \neg A \quad a : B \end{array} \qquad \begin{array}{c} a : \neg(A \rightarrow B) \\ | \\ a : A \\ a : \neg B \end{array}$$

Treffen wir auf einen Knoten $a : \Diamond A$, dann wissen wir, daß es einen Index b geben muß, der von a aus erreichbar ist, also Rab, und an dem A wahr ist. (Über diesen Index b dürfen wir nichts weiter annehmen; er darf also in der Konstruktion nicht zuvor schon vorkommen.) Im Fall $a : \Box A$ dürfen wir A an allen Indizes, die von a aus erreichbar sind, eintragen. Die Menge dieser Indizes notieren wir wie zuvor mit $R(a)$. Der Ausdruck $R(a) : A$ ist so etwas wie eine Aufforderung, A an einen Index weiterzugeben, sobald sich die Möglichkeit dazu bietet. Den Operator $\Diamond$ denken wir uns definiert ($\Diamond = \neg\Box\neg$), so daß wir nur diese Regeln brauchen:

Tbl (Mod)

$$\begin{array}{c} a : \Box A \\ | \\ R(a) : A \end{array} \qquad \begin{array}{c} a : \neg\Box A \\ | \\ Rab \ (b \text{ neu!}) \\ b : \neg A \end{array} \qquad \begin{array}{c} R(a) : A \\ Rab \\ | \\ b : A \end{array}$$

Die *Weitergaberegel* rechts ist so zu lesen: Wenn es auf einem Pfad einen Knoten mit $R(a) : A$ und einen mit Rab gibt, dann dürfen wir diesen Pfad um einen neuen Knoten $b : A$ verlängern.

Nun ist es an der Zeit für einige Definitionen.

- *Beschriftungen* sind Ausdrücke der Form $a : A$, $R(a) : A$ oder Rab.
- Ein *Tableau* ist ein Baum beschrifteter Knoten. Bäume haben genau einen Anfangsknoten, *Wurzel* genannt. Alle Knoten eines Baumes (außer der Wurzel) haben genau einen unmittelbaren Vorgängerknoten. Die Endknoten eines Baums nennen wir seine *Blätter*. Ein *Pfad* durch einen Baum ist eine Folge von Knoten, beginnend an einem Blatt und in gerader Linie endend an der Wurzel.
- Der Baum mit dem einen beschrifteteten Knoten $(a : A)$ ist ein *Tableau für die Formel* A. Wenn $\mathcal{T}$ ein Tableau für A ist und $\mathcal{T}'$ aus $\mathcal{T}$ durch Anwendung einer der Tableau-Regeln entsteht, dann ist auch $\mathcal{T}'$ ein Tableau für A.
- Ein Pfad in einem Tableau heißt *geschlossen*, wenn es darin Knoten $a : A$ und $a : \neg A$ gibt; anderenfalls ist der Pfad *offen*. (Unter geschlossene Pfade ziehen wir einen Strich.) Ein Tableau ist geschlossen, wenn alle Pfade darin geschlossen sind.
- Eine Formel ist genau dann *gültig* (im Tableau-Verfahren), wenn es ein geschlossenes Tableau für $\neg A$ gibt.

Aus einem offenen Pfad lassen sich Bedingungen für eine Interpretation I ablesen so, daß das Modell (W, R, I) eine gegebene Formel A erfüllt. Gibt es in einem Tableau für A keine offenen Pfade, dann ist A offenbar nicht erfüllbar. In diesem Fall muß $\neg A$ gültig sein im Rahmen (W, R). Da dieser Rahmen in aller Regel beliebig aus einer Rahmenklasse gewählt wurde, so haben wir auf diese Weise die Gültigkeit der Formel in der Rahmenklasse gezeigt. Die bisher genannten Regeln sehen keine Bedingungen für die Relation R vor. Also koinzidiert für diese Grundregeln Tableau-Gültigkeit mit Gültigkeit in $\mathbb{K}$ und, *a fortiori*, mit Beweisbarkeit in $\mathbf{K}$.

Im folgenden geben wir einige Beispiele an. Dabei wollen wir etwas abkürzen. Prüfen wir zum Beispiel eine Implikation $A \to B$, dann beginnen wir gleich mit der Wurzel $(a : A, a : \neg B)$. Manchmal fassen wir auch mehrere Regelanwendungen zusammen; dann generieren wir statt mehrerer Knoten nur einen. In der Praxis empfiehlt es sich, abgearbeitete Knoten als erledigte zu markieren.

Es erspart oft Arbeit, Verzweigungen aufzuschieben. Um Verzweigungen zu umgehen, können wir die folgenden abgeleiteten Regeln verwenden:

$$
\begin{array}{c}
a : A \vee B \\
a : \neg A \\
| \\
a : B
\end{array}
\qquad
\begin{array}{c}
a : \neg(A \wedge B) \\
a : \neg A \\
| \\
a : \neg B
\end{array}
\qquad
\begin{array}{c}
a : A \to B \\
a : A \\
| \\
a : B
\end{array}
\qquad
\begin{array}{c}
a : A \to B \\
a : \neg B \\
| \\
a : \neg A
\end{array}
$$

Mit diesen zusätzlichen Regeln werden Tableaus für die meisten Formeln, an denen wir ein Interesse haben können, linear.

Beispiel 1. $\Box(A \to B) \to (\Box A \to \Box B)$.

$$
\begin{array}{c}
a : \Box(A \to B) \\
a : \neg(\Box A \to \Box B) \\
| \\
a : \Box A \\
a : \neg\Box B \\
| \\
R(a) : A \to B \\
R(a) : A \\
| \\
Rab \\
b : \neg B \\
| \\
b : A \\
b : A \to B \\
| \\
\underline{b : B}
\end{array}
$$

In den folgenden Beispielen möge der Leser selbst die erläuternden Pfeile einfügen!

Beispiel 2. $\Box\Diamond A \to \Diamond\Box A$.

$$
\begin{array}{c}
a : \Box\Diamond A \\
a : \neg\Diamond\Box A \\
| \\
R(a) : \Diamond A \\
R(a) : \neg\Box A
\end{array}
$$

Weitere Regeln sind nicht anwendbar; der Pfad bleibt also offen: Nicht gültig in $\mathbb{K}$. Aber welches Gegenmodell zur Formel können wir hier ablesen? Jedes Modell mit nur einem Punkt dient dem Zweck! Sei a dieser singuläre Punkt. Dann ist jede Formel mit anfänglichem $\Box$ am Punkt a trivialerweise wahr, denn kein Punkt b erfüllt den Wenn-Teil der Wahrheitsbedingung "($\forall b$) wenn Rab, dann ..." . Das nächste Beispiel behandelt die Umkehrung.

Beispiel 3. $\Diamond\Box A \to \Box\Diamond A$.

$$
\begin{array}{c}
a : \Diamond\Box A \\
a : \neg\Box\Diamond A \\
| \\
Rab \\
b : \Box A \\
| \\
Rac \\
c : \neg\Diamond A \\
| \\
R(b) : A \\
R(c) : \neg A
\end{array}
$$

Wieder kein geschlossener Pfad. Dem Tableau ist unmittelbar ein Gegenmodell abzulesen: $(\{a, b, c, b', c'\}, Rab, Rac, , Rbb', Rcc', I)$ mit $I(P, b') = 1$ und $I(P, c') = 0$. Solche Gegenmodelle können wir auch übersichtlich so notieren:

$$
\begin{array}{cccccc}
 & a & b & c & b' & c' \\
A & * & * & * & 1 & 0
\end{array}
\quad Rab, Rac, Rbb', Rcc'.
$$

Die erste Zeile gibt die Modellmenge W an; die zweite Zeile stellte die nötige Bedingung für I auf; rechts daneben wird R festgelegt. Ein Stern $*$ deutet an, daß der Wahrheitswert von A (eigentlich: eines beliebigen für A einzusetzenden Atoms) am darüberstehenden Punkt gleichgültig ist. Das Gegenmodell läßt sich weiter vereinfachen. Wir brauchen nur ein Modell mit drei Punkten, Rab und Rac, wobei b und c R-Endpunkte sind. An solchen Endpunkten sind, wie bereits festgestellt, alle Formeln mit anfänglichem $\Box$ trivialerweise wahr. Also haben wir $b : \Box A$ und $c : \Box\neg A$. Aber dann auch $a : \Diamond\Box A$ und $a : \Diamond\Box\neg A\,(= \neg\Box\Diamond A)$:

$$
\begin{array}{cccc}
 & a & b & c \\
A & * & * & *
\end{array}
\quad Rab, Rac.
$$

Wenn wir nun die Rahmenklasse so einschränken, daß R *konvergent* ist, d.h.

$$Rab \text{ und } Rac \Rightarrow \exists d : Rbd \text{ und } Rcd,$$

dann können wir das Tableau so fortführen:

$$\begin{array}{c} \vdots \\ | \\ Rbd \\ Rcd \\ | \\ d : A \\ d : \neg A \\ \hline \end{array}$$

Beispiel 4. Auch das folgende Tableau für $\Box A \to \Box\Box A$ bleibt ohne weitere Annahmen über R offen:

$$\begin{array}{c} a : \Box A \\ a : \neg\Box\Box A \\ | \\ Rab \\ b : \neg\Box A \\ | \\ Rbc \\ c : \neg A \\ | \\ R(a) : A \end{array}$$

Die Bedingung, welche R erfüllen muß, um dem Tableau zum Abschluß zu verhelfen, ist unschwer auszumachen.

Beispiel 5. $\Diamond A \to (\Box(A \to \Box A) \to A)$.[34]

$$\begin{array}{c} a : \Diamond A \\ a : \Box(A \to \Box A) \\ a : \neg A \\ | \\ Rab \\ b : A \\ | \\ R(a) : A \to \Box A \\ | \\ b : A \to \Box A \\ | \\ b : \Box A \\ | \\ R(b) : A \end{array}$$

Wenn wir nun aus Rab auf Rba schließen könnten (Symmetrie!), dann würde aus $R(b) : A$ folgen, daß $a : A$, und der Pfad wäre geschlossen. Ohne Symmetrie zeigt das Tableau folgendes Gegenmodell an:

$$\begin{array}{ccc} & a & b \\ A & 0 & 1 \end{array} \quad Rab.$$

Beispiel 6. $\Diamond A \to \Box A$ unter der Bedingung, daß R funktional ist,

$$Rab \text{ und } Rac \Rightarrow b = c.$$

(Wir benennen diesmal den Wurzelindex mit b.)

$$\begin{array}{c} b : \Diamond A \\ b : \neg\Box A \\ | \\ Rbc \\ c : A \\ | \\ Rbd \\ d : \neg A \\ | \\ c = d \\ | \\ \underline{c : \neg A} \end{array}$$

[34] Die Formel drückt gewissermaßen den Kern des ontologischen Gottesbeweises aus. Sei A: Gott existiert.

Noch einmal Beispiel 2. $\Box\Diamond A \to \Diamond\Box A$. Wir wollen nun annehnmen, daß R funktional (s.o.) und erweiterbar ist, d.h. $\forall a \exists b : Rab$. Diese letzte Bedingung erlaubt es uns, das Tableau 2 um den Knoten Rab zu verlängern:

$$\begin{array}{c} \vdots \\ | \\ Rab \\ \vdots \end{array}$$

Sodann hängen wir das Tableau 6 an. Die Formel ist ein Theorem des Systems **KD**!.

Übungen.

a) $\Box\neg A \to \Box(A \to B)$.
b) $\Box(A \to B) \land \Box(A \to \neg B) \to \neg\Diamond A$.
c) $\neg\Diamond\Diamond\Box(A \lor B) \to \Box\Box\Diamond\neg A$.
d) $\Diamond(A \to \Box B) \to (\Box A \to \Diamond B)$ in reflexiven Rahmen.
e) $\Box\Diamond\Box\Diamond A \to \Box\Diamond A$ in reflexiven Rahmen.
f) $\Box(\Box A \to B) \to \Box(A \to \Box B)$.
g) $\Box(A \lor \Diamond B) \to \Box A \lor \Diamond B$.

Die Grundregeln (Boole) und (Mod) richten das Tableau-Verfahren so ein, daß genau die in $\mathbb{K}$ gültigen und also in **K** beweisbaren Formeln identifiziert werden können. Für Erweiterungen von **K** werden die jeweils korrespondierenden Rahmenbedingungen als weitere Tableau-Regeln formuliert; also etwa:

$$\mathit{Tbl}\ (R) \qquad \begin{array}{c} \\ | \\ Raa \end{array}\ \text{(Refl)} \qquad \begin{array}{c} Rab \\ Rbc \\ | \\ Rac \end{array}\ \text{(Trans)} \qquad \begin{array}{c} Rab \\ | \\ Rba \end{array}\ \text{(Sym)} \qquad \begin{array}{c} a : A \\ | \\ Rab \end{array}\ \text{(Erw)}$$

Tableau-Verfahren stellen gewissermaßen die "Syntaktisierung" der modalen Modelltheorie dar. Es handelt sich im Prinzip um ein rein syntaktisches Beweisverfahren, welches Ausdrücke der Form $a : A$, Rab, $a = b$ etc. erzeugt und diese in einer Baumstruktur anordnet. Diese Ausdrücke sind so unmittelbar auf die Semantik einer modalen Sprache bezogen, daß sich daraus recht trivial die Beobachtung ergibt, daß das Verfahren die in einer bestimmten Rahmenklasse gültigen Formeln richtig und vollständig identifiziert.

Wenn man mit Tableaus arbeitet, dann stellt sich schnell der Eindruck ein, daß diese nicht nur ein bequemes Mittel sind sich von der Erfüllbarkeit oder Unerfüllbarkeit einer Formel zu überzeugen. Das Verfahren hat etwas mechanisches an sich. Die Tableaus werden durch die Regeln in eindeutiger Weise vorangetrieben. Nur die Reihenfolge der Regelanwendung ist nicht bestimmt. Aber darauf kommt es gar nicht an. Die Knoten können in beliebiger Reihenfolge abgearbeitet werden. (Wenn wir es so möchten, können wir durch zusätzliche Vereinbarungen eine bestimmte Reihenfolge erzwingen, so daß diese Wahlmöglichkeit abgeschnitten wird.) Mehr noch: Wenn wir einmal von der Weitergabe-Regel absehen, dann haben alle Regeln die Eigenschaft die Komplexität der jeweils bearbeiteten Formel zu reduzieren. Da jede Formel eine endliche Komplexität hat, drängt sich die Vermutung auf, daß das modale Tableaus-Verfahren ein *Entscheidungsverfahren* für die Eigenschaft der Gültigkeit in einer Rahmenklasse darstellt.

Diese Vermutung trifft nur zum Teil zu. Die Weitergabe-Regel kann im Verein mit bestimmten R-Regeln unendliche Tableaus erzeugen. Betrachten wir zum Beispiel die Kombination von (Trans) und (Erw). (Das Beispiel (Trans)+(Refl) ist ähnlich.) Wir beginnen ein Tableau mit $(a : \Box A)$ und setzen dann so fort:

$$\begin{array}{c} a : \Box A \\ | \\ R(a) : A \end{array}$$

Nun erlaubt (Erw) die Fortsetzung $\begin{array}{c} | \\ Rab \end{array}$. Daraufhin geben wir A weiter an b: $\begin{array}{c} | \\ b : A \end{array}$. (Erw), gefolgt von (Trans) und Weitergabe führt nun zu

$$\begin{array}{c} | \\ Rbc \\ | \\ Rac \\ | \\ c : A \\ \vdots \end{array}$$

Die Ellipse deutet an, daß es mit diesen Regeln nun unendlich so weitergeht. Es können stets neue Knoten erzeugt werden, an welche die Formel A weitergegeben wird. Das muß in einem gegebenen Fall nicht dazu führen, daß das Tableau nicht schließt. Aber ohne weitere Vorkehrungen schützt uns im allgemeinen nichts davor, daß **KD4**-Tableaus unendlich fortgesetzt werden können und somit die Frage offen bleibt, ob ein Pfad geschlossen oder offen ist. Tatsächlich können wir das durch eine zulässige Regel verhindern, welche redundante Anwendungen der Weitergabe-Regel verbietet. Das so modifizierte Tableaus-Verfahren ist dann ein Eintscheidungsverfahren auch für **KD4**. Wir wollen die Frage der Enscheidbarkeit nun jedoch allgemeiner mit anderen Mitteln angehen.

13. Filtrierung und Entscheidbarkeit

Ein Modell enthält oft mehr Information, als es für die Beurteilung einer einzelnen Formel A an einem Punkte a nötig ist. In diesem Fall können wir aus a ein Teilmodell $\mathcal{M}^a$ erzeugen, welches nur die mit a verbundenen Punkte (im Sinne von R) enthält. Aber auch dieses reduzierte Modell enthält noch redundante Information. Denn die Interpretationsfunktion I des Modells weist an jedem Punkte jedem Atom einen Wahrheitswert zu. Für die Beurteilung von A interessiert aber nur, welche Werte die Atome, die in A vorkommen, haben.

Gleiches gilt für das kanonische Modell einer Logik **L**. Für die Beurteilung einer einzelnen Formel A sind alle Punkte, die sich nur im Hinblick auf Formeln unterscheiden, die in A gar nicht vorkommen, einerlei. Anders ausgedrückt: Punkte, die im Hinblick auf die Teilformeln von A übereinstimmen, können wir, ohne relevante Information zu verlieren, identifizieren. Bildlich vorgestellt, "filtern" wir das Modell durch die Formel A und erhalten als Filtrat ein deutlich kleineres Modell (auf einem deutlich kleineren Rahmen), welches alles vom Original-Modell enthält, was für die Beurteilung von A wichtig ist.

Das Bild berechtigt zu einer Hoffnung. Jede Formel hat nur endlich viele Teilformeln. Wenn wir uns fragen, ob eine Formel A gültig ist in einer Klassen von Rahmen, dann können wir vielleicht die Menge der Modelle auf diesen Rahmen durch A filtrieren und erhalten so nicht nur eine Menge endlicher Modelle, sondern eine endliche Menge endlicher Modelle, in denen wir nun A prüfen müssen. Das wäre dann eine endliche Aufgabe, was bedeuten würde: Von jeder Formel ließe sich in endlicher Weise entscheiden, ob diese gültig ist in der Rahmenklasse. Und wenn wir zudem wissen, daß eine bestimmte, axiomatisch formulierte Logik **L** genau diese Rahmenklasse beschreibt, dann hätten wir auch ein Entscheidungsverfahren für die Eigenschaft der Ableitbarkeit in **L**.

Die Entscheidbarkeit einer Logik kann unter zwei Aspekten betrachtet werden:

· Wir können danach fragen, ob eine Logik überhaupt entscheidbar ist.
· Wir können an einem handhabbaren Verfahren interessiert sein, welches die Frage nach der Gültigkeit bzw. Ableitbarkeit einer Formel möglichst unaufwendig beantwortet.

Der letzte Aspekt stand im vorhergehenden Abschnitt im Mittelpunkt. Daß das dort vorgeführte Verfahren für einige Modallogiken recht gut zu funktionieren scheint, garantiert aber nicht, daß es sich für jede Logik adaptieren läßt. Vorgängig ist in jedem Fall die Frage, ob eine Logik überhaupt entscheidbar ist. Die Art der Antwort, die wir hier geben werden, fußt jeweils auf der Angabe eines Entscheidungsverfahrens. Jedoch sind diese Verfahren im Sinne des zweiten Aspektes selten brauchbar. Zwar werden wir z.B. im Rahmen des Filtrierungsverfahrens einer berechenbaren Kardinalitätsschranke für Modelle begegnen, welche die Endlichkeit des Verfahrens garantiert; für praktische Zwecke liegt diese Schranke jedoch oft zu hoch. Über die Komplexität des Entscheidungsproblems in Modallogiken informiert [31, Kap. 6].

Es gibt verschiedene Ansätze, das Entscheidungsproblem anzugehen. Wir benutzen hier die Technik der Filtrierung – in einem präzisen Sinne, den wir gleich erklären werden und der auf Arbeiten von Lemmon, Scott und Segerberg zurückgeht. Andere Techniken werden in [43] skizziert und in [31] *in extenso* vorgestellt. Praktische sinnvolle Entscheidungsverfahren findet der Leser auch in [141] und [227]. Diese mögen auf den ersten Blick etwas anders aussehen als unsere Tableaus, die Idee ist jedoch dieselbe: Es geht um die systematische Suche nach kleinen Gegenmodellen.

Filtrierung. Wenn A eine Formel ist, dann sei $\mathrm{tf}(A)$ die Menge ihrer Teilformeln. (Die Definition läßt sich natürlich erweitern für Formelmengen X: $\mathrm{tf}(X) = \bigcup\{\mathrm{tf}(A) : A \in X\}$.)

Es sei $\mathcal{M} = (W, R, I)$ ein Modell und A eine Formel. Für jeden Punkt $a \in W$ sei

$$A_a := \{B \in \mathrm{tf}(A) : a \models B\}$$

die Menge der Teilformeln von A, die am Punkt a wahr sind. Ferner definieren wir

$$a \sim_A b \text{ gdw } A_a = A_b$$

und

$$[a]_A := \{b \in W : a \sim_A b\}, \text{ sowie } \; W_A := \{[a]_A : a \in W\}.$$

Da $\sim_A$ eine Äquivalenzrelation und also $[a]_A$ eine Äquivalenzklasse ist, folgt unmittelbar aus den Definitionen, daß

$$(*) \qquad [a]_A = [b]_A \text{ gdw } A_a = A_b.$$

Nun nehmen wir die Menge W_A selbst als Trägermenge eines Modells, welches wir so definieren:

Definition 62. Ein Modell $\mathcal{M}_A = (W_A, R_A, I_A)$ ist ein *Filtrat* von $\mathcal{M} = (W, R, I)$ durch die Formel A (kurz A-Filtrat), wenn die folgenden Bedingungen erfüllt sind:

1. W_A ist, wie oben, die Menge der Äquivalenzklassen unter $\sim_A$.
2. $R_A[a]_A[b]_A$ gdw $\exists x \in [a]_A$ und $\exists y \in [b]_A : Rxy$.
3. $I_A(P, [a]_A) = 1$ gdw $I(P, a) = 1$, für alle Atome $P \in \mathrm{tf}(A)$.

Lemma 63. *Für jede Formel A mit n Teilformeln ist W_A endlich und hat höchstens 2^n Elemente.*

Beweis. Wir definieren eine Abbildung $f : W_A \longrightarrow \wp(\mathrm{tf}(A))$, die jeder Punktemenge in W_A die Menge der Teilformeln von A zuordnet, die an den Punkten der Menge wahr sind:

$$f([a]_A) = A_a.$$

Aufgrund von $(*)$ wissen wir, erstens, daß die so definierte Funktion f unabhängig davon ist, wie wir die Punktemengen in W_A repräsentieren; denn wenn $[a]_A = [b]_A$, dann $f([a]_A) = f([b]_A)$. Ferner wissen wir, daß die Funktion injektiv ist; denn wenn $f([a]_A) = f([b]_A)$, dann $[a]_A = [b]_A$. Da also verschiedene Argumente verschiedene Werte bekommen, kann die Kardinalität des Argumentbereichs W_A nicht größer sein als die des Wertebereichs, $|\wp(\mathrm{tf}(A))| = 2^n$ (mit $|\mathrm{tf}(A)| = n$ nach Voraussetzung). ■

Nun ist es so, daß wir für den Beweis des Filtrierungssatzes 64 tatsächlich nur auf Eigenschaften der Relation R_A zurückgreifen, die schwächer sind als die Bedingung 2 in der obigen Definition. Also können wir für diesen Satz den Begriff eines Filtrats etwas allgemeiner definieren, indem wir die Bedingung 2 durch die folgende ersetzen:

2* a. Wenn Rab, dann $R_A[a]_A[b]_A$, und
b. wenn $R_A[a]_A[b]_A$, dann $\forall C$: falls $\Box C \in \mathrm{tf}(A)$ und $a \models \Box C$, dann $b \models C$.

Die engere Bedingung 2 definiert das kleinste Filtrat, während die schwächere Bedingung 2* auch größere Filtrate zuläßt. (Der Nachweis, daß aus 2 die Bedingung 2* folgt, ist eine einfache Übung.)

Satz 64. (Filtrierung) *Es sei $\mathcal{M}_A$ ein A-Filtrat eines Modells $\mathcal{M}$. Dann gilt für alle $B \in \mathrm{tf}(A)$:*

$$a \models B \text{ in } \mathcal{M} \text{ gdw } [a]_A \models B \text{ in } \mathcal{M}_A.$$

Beweis. Induktion über den Aufbau der Teilformeln B von A.
Fall $B = P$: Gegeben durch Bedingung 3.
Fälle $B = \neg C$ und $B = C \wedge D$: Unmittelbar aus der Induktionshypothese.
Fall $B = \Box C$: (LR) folgt aus Bedingung 2*b, (RL) aus 2*a. ■

DEFINITION 65. Eine Logik $\mathbf{L}$ hat die *endliche Rahmeneigenschaft*, wenn es für jede Formel $A \notin \mathbf{L}$ einen endlichen Rahmen gibt mit $\mathcal{R} \models \mathbf{L}$ und $\mathcal{R} \not\models A$.

Es ist leicht zu sehen, daß $\mathbf{L}$ genau dann die endliche Rahmeneigenschaft hat, wenn $\mathbf{L}$ mit der Menge der in den *endlichen* Rahmen für $\mathbf{L}$ gültigen Formeln übereinstimmt, d.h.

$$\mathbf{L} = Th(Endl(Rm(\mathbf{L}))).$$

SATZ 66. *$\mathbf{K}$ hat die endliche Rahmeneigenschaft.*

BEWEIS. Angenommen $A \notin \mathbf{K}$. Da $\mathbf{K} = Th(\mathbb{K})$, so gibt es einen Punkt a in einem Modell $\mathcal{M}$ auf einem Rahmen $\mathcal{R}$ in $\mathbb{K}$ so, daß $a \not\models A$. Wir konstruieren nun ein A-Filtrat $\mathcal{M}_A = (W_A, R_A, I_A)$ von $\mathcal{M}$. Nach dem Filtrierungssatz 64 haben wir $[a]_A \not\models A$ in $\mathcal{M}_A$ und also $\mathcal{R}_A \not\models A$, wobei $\mathcal{R}_A$ nach Lemma 63 endlich ist. Ferner sind alle Theoreme von $\mathbf{K}$ wie in jedem Kripke-Rahmen, so auch in $\mathcal{R}_A$ gültig, d.h. $\mathcal{R}_A \models \mathbf{K}$. ∎

Aus dem Satz folgt ein weiteres Vollständigkeitsresultat für das Basissystem $\mathbf{K}$: $\mathbf{K} = Th(Endl(\mathbb{K}))$, d.h. $\mathbf{K}$ ist determiniert durch die Klasse aller endlichen Kripke-Rahmen – ein weiterer Beweis übrigens (nach Satz 33), daß die Bedingung (Endl) nicht modal definierbar ist. Ferner ergibt sich im Verein mit Lemma 63 eine wichtige Information über die Größe solcher Modelle, die wir betrachten müssen, um herauszufinden, ob eine gegebene Formel ein Theorem von $\mathbf{K}$ ist. (Von dieser Information werden wir auf p. 171 Gebrauch machen.)

KOROLLAR 67. *Es sei $\mathbb{K}^m$ ($m \in \mathbf{N}$) die Klasse aller Rahmen mit höchstens m Punkten. Wenn eine Formel A n Teilformeln hat, dann $A \in \mathbf{K}$ gdw $A \in Th(\mathbb{K}^{2^n})$.*

BEWEIS. In der einen Richtung genügt die Beobachtung $\mathbf{K} \subseteq Th(\mathbb{K}) \subseteq Th(\mathbb{K}^{2n})$. Die andere Richtung folgt aus Lemma 63 und Satz 66. ∎

SATZ 68. *Die Logiken $\mathbf{T}$, $\mathbf{B}$, $\mathbf{S4}$, und $\mathbf{S5}$ haben die endliche Rahmeneigenschaft.*

BEWEIS. Wir bauen auf dem Beweis von Satz 64 auf, betrachten jedoch jeweils das kanonische Modell und bilden daraus ein Filtrat. Die Rahmenkomponente des kanonischen Modells erfüllt die jeweiligen korrespondierenden Bedingungen und ist charakteristisch für die betrachtete Logik. Es ist nun zu zeigen, daß auch das Filtrat die Bedingungen erfüllt. Für $\mathbf{T}$ (Reflexivität) und $\mathbf{B}$ (Reflexivität und Symmetrie) genügt dazu der Rückgriff auf Bedingung 2 der Definition 62.

Im Falle von **S4** und **S5** ist es nicht so, daß jedes Filtrat eines transitiven Rahmens selbst transitiv ist. Wir müssen also passende Filtrate erst definieren. Wenn (W, R, I) das jeweilige kanonische Modell ist, dann sei W_A wie zuvor und für **S4** definieren wir:

$$R_A[a]_A[b]_A \text{ gdw } \forall \Box B \in \mathrm{tf}(A) : a \models \Box B \Rightarrow b \models \Box B.$$

Im Falle von **S5** definieren wir:

$$R_A[a]_A[b]_A \text{ gdw } \forall \Box B \in \mathrm{tf}(A) : a \models \Box B \iff b \models \Box B.$$

Nun läßt sich jeweils zeigen, daß (W_A, R_A) Filtrate (im weiteren Sinne 2*) sind, und daß die Eigenschaften der Relation (Transitivität bzw. Äquivalenz) sich vom kanonischen Rahmen auf sein Filtrat übertragen. (Details z.B. in [81, pp 78f.].) ■

Man beachte, daß aus der endlichen Rahmeneigenschaft für eine Logik **L** folgt, daß es für jede Formel $A \notin L$ ein endliches Modell $\mathcal{M}$ für **L** gibt mit $\mathcal{M} \not\models A$; d.h. die endliche Rahmeneigenschaft impliziert die *endliche Modelleigenschaft.* (Dazu sehe man noch einmal den Beweis von Satz 66 durch. Die Verallgemeinerung liegt dann auf der Hand.) So erhalten wir das weitere

KOROLLAR 69. *Die Logiken* **K**, **T**, **B**, **S4**, *und* **S5** *haben die endliche Modelleigenschaft.*

Tatsächlich koinzidieren endliche Modell- und Rahmeneigenschaft für normale modale Logiken; d.h. es gilt:

LEMMA 70. *Eine normale Modallogik hat genau dann die endliche Rahmeneigenschaft, wenn sie die endliche Modelleigenschaft hat.*

BEWEIS. Die eine Richtung liegt, wie gesagt, auf der Hand. Für die andere Richtung nehmen wir an, daß (1) **L** die endliche Modelleigenschaft hat, und daß (2) $A \notin$ **L**. (Wir weisen einen endlichen Rahmen für **L** vor, der A widerlegt.)

Wir beginnen mit einer allgemeinen Beobachtung über Filtrate. In Filtraten sind alle Punkte durch Formeln voneinander unterschieden ("separiert"):

(A) Wenn $\mathcal{M}_F$ ein Filtrat ist (F eine Formel oder Formelmenge), dann gilt für alle distinkten Punkte x, y in W_F, $\exists B$: $x \models B$ und $y \not\models B$.

(Vgl. die Def. auf p. 165.) Denn wenn $x = [a]_F$ und $y = [b]_F$ (mit a, b in $\mathcal{M}$) distinkt sind, dann $F_a \neq F_b$. Also gibt es ein $B \in \mathrm{tf}(F)$ mit $a \models B$ und $b \not\models B$.

Aus (1–2) folgt, daß es ein endliches Modell $\mathcal{M}$ gibt mit $\mathcal{M} \models \mathbf{L}$ und $\mathcal{M}, a \not\models A$ für einen Punkt a in $\mathcal{M}$. Wir betrachten nun das Filtrat $\mathcal{M}_X = (W_X, R_X, I_X)$ mit dem Filter $X = \mathbf{L} \cup \{A\}$. Da $\mathcal{M}_X$ endlich ist, so ist auch $\mathcal{R}_X$ endlich. Nach Satz 64 (Filtrierung) gilt $\mathcal{M}_X, [a]_X \not\models A$, also $\mathcal{R}_X \not\models A$. Da $\mathcal{M} \models \mathbf{L}$ so gilt ferner aufgrund des Filtrierungssatzes, daß $\mathcal{M}_X \models \mathbf{L}$. Es bleibt zu zeigen, daß $\mathcal{R}_X \models \mathbf{L}$. Da nach (A) die Punkte in $\mathcal{R}_X$ separiert sind, folgt die Konklusion unmittelbar aus dieser Beobachtung:

(B) Wenn $\mathcal{M} = (\mathcal{R}, I)$ ein Modell mit separierten Punkten ist und $\mathcal{M} \models A$, dann $\mathcal{R} \models A$ (für beliebige Formeln A).

Der Beweis von (B) sei hier nur skizziert; Details findet der Leser z.B. in [81, pp. 79ff.] und [31, pp. 146ff.].

Wir beginnen mit der Beobachtung, daß in separierten Modellen jeder Punkt x eine charakteristische Formel hat:

(a) Für alle Punkte $x \in W$ gibt es eine Formel A so, daß $\forall y \in W$, $y \models A$ gdw $x = y$.

In separierten Modellen lassen sich auch beliebige Teilmengen von W durch Formeln charakterisieren:

(b) Für alle $U \subseteq W$, $\exists A$: $\forall x \subseteq W$, $x \models A$ gdw $x \in U$.

Interpretationen unterscheiden sich dadurch, daß sie Atome auf verschiedene Teilmengen von W abbilden. Eine Interpretation I' nennen wir daher (in einem Modell) *definierbar*, wenn jede der Teilmengen $I'_P = \{x \in W : I'(P, x) = 1\}$ die Bedingung (b) instantiiert. Das ist nach (b) für separierte Modelle der Fall. Sei $\mathcal{M} = (\mathcal{R}, I)$ ein solches Modell. Dann folgt:

(c) $\mathcal{M}$ kann alle Interpretationsvarianten I' auf $\mathcal{R}$ definieren.

Nun gilt aber

(d) Wenn $\mathcal{R}, I \models A$, dann für alle in diesem Modell definierbaren Interpretationsvarianten $(\mathcal{R}, I')$: $\mathcal{R}, I' \models A$.

Aus (c) und (d) folgt, daß $\mathcal{R} \models A$, wie gewünscht. ■

Endliche Rahmeneigenschaft und endliche Modelleigenschaft sind also für normale Modallogiken austauschbare Begriffe. Im Rest dieses Kapitels werden wir dem letzteren den Vorzug geben.

Entscheidbarkeit. Eine Logik ist entscheidbar, wenn es ein Verfahren gibt, mit dem wir für jede Formel A nach endlich vielen Schritten feststellen können, ob sie ein Theorem der Logik ist, d.h. $A \in \mathbf{L}$ oder $A \notin \mathbf{L}$.

A. Angenommen eine Logik $\mathbf{L}$ liegt endlich axiomatisiert vor (so, wie wir einige normale Modallogiken oben kennengelernt haben), d.h.

$$\mathbf{L} = \mathbf{K} + S_1 + \cdots + S_n,$$

wobei jedes S_i ein modales Schema ist. Wenn eine Formel A ein Theorem ist, dann gibt es ein Verfahren, das nach endlichen Schritten die Frage ob A ein Theorem ist, beantwortet:

1. Wir wählen eine Abzählung aller Ableitungen in dem axiomatischen System für **L**.
2. Wir prüfen für jede Ableitung, ob sie terminal A enthält.
3. Irgendwann, d.h. in einem Schritt $k \in \mathbf{N}$, werden wir auf eine solche Ableitung treffen.

Wenn A jedoch kein Theorem ist, dann wird uns dieses Verfahren das nicht im allgemeinen nach endlich vielen Schritten bestätigen können. Denn in diesem Fall müßten wir uns davon überzeugen, daß es keine Ableitung mit terminalem A gibt; d.h. wir müßten alle Ableitungen prüfen – und davon gibt es unendlich viele. Die Frage, ob A ein Theorem sei, ist mit diesem Verfahren also nur "zur Hälfte" entscheidbar, d.h. nur im Falle, daß A ein Theorem ist.

B. Angenommen **L** hat die endliche Modelleigenschaft, d.h. für jedes Nicht-Theorem gibt es ein endliches widerlegendes Modell für **L**. Wenn A kein Theorem ist, dann können wir so vorgehen:

1. Wir wählen eine Abzählung aller endlichen Modelle für **L**.
2. Die Prüfung von A in einem solchen Modell ist eine endliche Angelegenheit.
3. Aufgrund der endlichen Modelleigenschaft von **L** werden wir schließlich in einem Schritt $k \in \mathbf{N}$ auf ein Modell stoßen, welches A widerlegt.

Offensichtlich ist die Frage, ob A ein Theorem sei, auch mit diesem Verfahren nur zur Hälfte entscheidbar, nämlich nur im Falle, daß A keines ist. Die zwei Hälften ergänzen sich jedoch zu einem Ganzen, indem wir beide Verfahren kombinieren. Daher:

Satz 71. *Jede endlich axiomatisierbare normale Modallogik mit der endlichen Modelleigenschaft ist entscheidbar.*

C. Im Falle von **K** ist es jedoch so:

1. Jede Formel A hat eine endliche Anzahl n von Teilformeln.
2. Also betrachten wir nach Satz 67 nur die endlichen Rahmen in $\mathbb{K}^{2^n}$.
3. Von diesen gibt es nur endlich viele. Denn in einem Rahmen (W, R) ist R eine Menge in $\wp(W^2)$ (manchmal auch $2^{(W^2)}$ geschrieben). Also gibt es $2^{(m^2)}$ verschiedene Relationen auf einer Menge W mit m Punkten. Daher gibt es für jedes $m \leq 2^n$ nur endlich viele Rahmen mit m

Punkten in W, die sich im Hinblick auf die Relation R unterscheiden können. Also ist $\mathbb{K}^{2^n}$ eine endliche Menge von Rahmen.
4. Für jeden Rahmen $\mathcal{R}$ in dieser Menge, genügt es (nach Satz 64), den Status von A in solchen Modellen zu prüfen, welche A-Filtrate darstellen.
5. Da Filtrate endlich sind, genügt es also, den Status von A an endlich vielen Punkten in endlich vielen Modellen auf einer endlichen Anzahl von Rahmen zu prüfen.

Wir haben ein Entscheidungsverfahren für **K**! Die "entscheidende" Information über **K** haben wir dem Satz 67 entnommen: Die endlichen Rahmen für **K**, die ein gegebenes Nicht-Theorem A falsifizieren, gehören zu einer Klasse $\mathbb{R}$ von Rahmen, die eine maximale Größe haben, welche sich jeweils aus der Anzahl der Teilformeln von A errechnen läßt. Daraus konnten wir auf die Endlichkeit der zu betrachtenden Rahmenklasse schließen. Diese Eigenschaft der Logik **K** wollen wir verallgemeinern für beliebige normale Modallogiken **L**.

Zunächst definieren wir für jede Formel A die Klasse $\mathbb{M}^{\mathrm{tf}}$ der Modelle für **L** welche A widerlegen und aus höchstens $2^{|\,\mathrm{tf}(A)|}$ Punkten bestehen:

$$\mathcal{M} \in \mathbb{M}^{\mathrm{tf}}(\mathbf{L}, A) \text{ gdw } \mathcal{M} \models \mathbf{L},\ \mathcal{M} \not\models A, \text{ und } |\mathcal{M}| \leq 2^{|\,\mathrm{tf}(A)|}.$$

Definition 72. Eine Logik **L** hat die *starke endliche Modelleigenschaft*, wenn es für jede Formel $A \notin \mathbf{L}$ ein Modell $\mathcal{M} \in \mathbb{M}^{\mathrm{tf}}(\mathbf{L}, A)$ gibt.

Im Falle von **K** ist es einfach zu entscheiden, ob ein gegebenes Modell zu $\mathbb{M}^{\mathrm{tf}}(\mathbf{K}, A)$ gehört. Wir genererieren einfach alle Modelle bis zur Größe $2^{|\,\mathrm{tf}(A)|}$ mit einem A-falsifiziendem Punkt (eine endliche Aufgabe). Diese werden immer Modelle für **K** sein. Wenn wir jedoch über **K** hinausgehen, dann ist sicher nicht jedes größenbegrenzte Modelle eines für die jeweils betrachtete Logik **L**. Damit das oben skizzierte Entscheidungsverfahren für eine solche Logik **L** funktioniert, müssen wir also in der Lage sein zu entscheiden, ob ein gegebenes Modell zu $\mathbb{M}^{\mathrm{tf}}(\mathbf{L}, A)$ gehört oder nicht.

Satz 73. *Eine Logik* **L** *ist entscheidbar, wenn* **L** *die starke endliche Modelleigenschaft hat und für jede Formel A, die Modellklasse* $\mathbb{M}^{\mathrm{tf}}(\mathbf{L}, A)$ *entscheidbar ist.*

Satz 74. *Die Logiken* **K**, **T**, **B**, **S4**, *und* **S5** *haben die starke endliche Modelleigenschaft.*

Beweis. Für **K** wurde die Eigenschaft schon bewiesen. Wir verfolgen den Nachweis der endlichen Modelleigenschaft für **T**, **B**, **S4**, und **S5** (Korollar 69) zurück. Wir stellen fest, daß die dazu konstruierten endlichen

Modelle Filtrate sind, und also (Lemma 63) maximal $2^{\mathrm{tf}(A)}$ enthalten, für jede zu prüfende Formel A. Also haben diese Logiken die *starke* endliche Modelleigenschaft.

SATZ 75. *Die Logiken* **K**, **T**, **B**, **S4**, *und* **S5** *sind entscheidbar.*

BEWEIS. Nach den vorhergehenden Sätzen bleibt zu zeigen, daß für jede der Logiken **L** (und für jede Formel A) die Klasse $\mathbb{M}^{\mathrm{tf}}(\mathbf{L}, A)$ entscheidbar ist. Jede dieser Logiken entsteht aus **K** durch Hinzufügung einer endlichen Anzahl von Schemata; z.B. $\mathbf{T} = \mathbf{K}{+}\mathrm{T}$, $\mathbf{S4} = \mathbf{K}{+}\mathrm{T} + 4$. Jedes der Schemata korrespondiert mit einer Bedingung erster Stufe auf Modellen. Jedes Modell kann in endlicher Weise darauf geprüft werden, ob es eine solche Bedingung erfüllt. Also ist jede der Klassen $\mathbb{M}^{\mathrm{tf}}(\mathbf{L}, A)$ entscheidbar. ∎

Aus dem Beweis des letzten Satzes folgt unmittelbar:

KOROLLAR 76. *Jede endlich axiomatisierbare Logik* **L** *mit der starken endlichen Modelleigenschaft ist entscheidbar.*

Dieses Korollar folgt unmittelbar schon aus Satz 71, und mit Satz 68 auch die Entscheidbarkeit aller bisher betrachteten axiomatischen Erweiterungen von **K** für die wir Vollständigkeitsbeweise führen konnten. Das sollte jedoch nicht den Blick verstellen auf den Umstand, daß Satz 75 hier nicht auf der teils beweistheoretischen Basis der Verfahren A und B, sondern allein auf der Basis des modelltheoretischen Verfahrens C gewonnen wurde. Zwar haben wir im Beweis des Satz 75 auf eine axiomatische Darstellung der jeweiligen Logik zurückgegriffen. Aber genausogut – und direkter – hätten wir die Logiken im Satz als modelltheoretisch gegeben betrachten können, d.h. als Mengen von Formeln, die in allen Modellen mit bestimmten Eigenschaften erster Stufe gültig sind.

Wie Chagrov und Zakhariaschev [50, p. 495] bemerken, sind die "meisten" Modallogiken unentscheidbar. Denn den unabzählbar vielen Modallogiken stehen nur abzählbare viele effektive Verfahren gegenüber. Das ist jedoch nicht weiter beunruhigend, denn die meisten "interessanten" Modallogiken sind endlich axiomatisierbare Erweiterungen der Basislogik **K** und als solche entscheidbar. Diese Beobachtung bringt uns schließlich zu einer hoffnungsvollen Einsicht, die über die eigentliche Modallogik hinausweist. Wir haben oben skizziert (siehe p. 114), in welchem Sinne Modallogik sich über den Weg einer Übersetzung als Notationsvariante der Prädikatenlogik erster Stufe darstellen läßt. So gibt es eine einfache und effektive Übersetzung t von modalen zu prädikatenlogischen Formeln. Das Bild $t(\mathcal{L}^{\Box})$ einer aussagenlogischen modalen Sprache $\mathcal{L}^{\Box}$ stellt so ein Fragment einer prädikatenlogischen Sprache $\mathcal{L}^1$ dar. Wenn wir dieses Fragment unter prädikatenlogischer Äquivalenz abschließen, dann erhalten wir die Menge $t^*(\mathcal{L}^{\Box})$

der Formeln in $\mathcal{L}^1$, die sich modal definieren lassen, das "modale Fragment" von $\mathcal{L}^1$. Das eröffnet die Perspektive, Teile der Prädikatenlogik mit den einfacheren Mitteln der modalen Aussagenlogik zu erkunden. Wenn **M** z.B. eine entscheidbare Modallogik ist (d.h, die Theoremmenge **M** in $\mathcal{L}^\square$ entscheidbar ist), dann ist auch die Menge $t(\mathbf{M})$ prädikatenlogischer Formeln entscheidbar. Dabei sind die aussagenlogischen Entscheidungsverfahren für die Modallogik oft deutlich effizienter als die prädikatenlogischen Verfahren für das entsprechende Fragment. Diese Perspektive wird in sogenannten *Beschreibungslogiken* systematisch entwickelt und erweitert.

Wir können für modale Fragmente der Prädikatenlogik auch modallogisch Schlüsselresultate für die Prädikatenlogik erbringen. So haben wir für "alle interessanten" Modallogiken **M** die endliche Modelleigenschaft nachgewiesen. Diese besagt – in etwas anderen Worten – daß jede erfüllbare Formel ein endliches Modell (für **M**) hat. Das können wir in die Prädikatenlogik so übertragen: Jede Formel im modalen Fragment von $\mathcal{L}^1$, welche in einem Modell erfüllbar ist, das die Bedingungen $t(\mathbf{M})$ erfüllt, ist in einem endlichen (und also abzählbaren) Modell (für $t(\mathbf{M})$) erfüllbar – eine Verstärkung des Löwenheim-Skolem Theorems für Fragmente der Prädikatenlogik.

14. Intuitionistische Logik als Modallogik

Intuitionistische Logik (**IL**), manchmal auch "Konstruktive Logik" genannt,[35] ist eine Logik, die schwächer ist als die klassische. Die Notwendigkeit einer solcher Abschwächung wird von verschiedenen Vertretern intuitionistischer Logik sehr verschieden begründet. Es scheint jedoch eine gemeinsame Kernidee zu sein, daß Wahrheit in irgendeiner Weise unter epistemischen Bedingungen stehen muß: Eine Aussage A wahr zu nennen, erfordert, im Hinblick auf A "ausreichend epistemisch positioniert" zu sein. Was immer das im Detail heißen mag, Intuitionisten sind sich darin einig, daß es nicht für jede Aussage A so ist, daß entweder A oder $\neg A$ wahr ist, d.h. daß man immer stark genug positioniert ist, entweder A oder $\neg A$ behaupten zu können. Also drückt für Intuitionisten das Schema

TND. $$A \vee \neg A$$

(*tertium non datur*: ein Drittes gibt es nicht) keine logische Wahrheit aus. TND ist das Prinzip, daß von einer Aussage oder deren Negation immer eines wahr ist. Das ist zu unterscheiden vom Prinzip der *Zweiwertigkeit*:

[35] So in der von Paul Lorenzen und Wilhelm Kamlah gegründeten Erlanger Schule. Außerhalb dieser Schule hat "Konstruktivismus" in der Logik und Grundlegung der Mathematik jedoch eine weitere Bedeutung, weshalb "Konstruktive Logik" keine glückliche Prägung ist.

Jede Aussage ist entweder wahr oder nicht wahr. Dieses Prinzip akzeptieren auch Intuitionisten. Sie verstehen zwar unter Wahrheit eine epistemisch bedingte Eigenschaft, aber diese Eigenschaft trifft in jedem Fall entweder zu oder nicht.

Nun kann man aus der klassischen Logik nicht einfach ein einzelnes Schema herausnehmen ohne auch an anderen Stellen einzugreifen. Die Schemata sind durch Ableitbarkeits- bzw. Folgerungsbeziehungen miteinander verknüpft. So haben wir klassisch

(1) $$(\neg A \vee A) \;\leftrightarrow\; \neg(A \wedge \neg A)$$

Das ist eine Instanz eines der DeMorgan'schen Gesetze. Wenn nun Intuitionisten die linke Seite der Äquivalenz ablehnen, heißt das, daß sie auch die rechte Seite, d.h. das Prinzip vom ausgeschlossenen Widerspruch ablehnen? Nein, $\neg(\neg A \wedge A)$ ist ein Theorem von **IL**. Intuitionisten unterbrechen daher die Verbindung (1) zwischen diesem Theorem und TND, indem sie das Schema

(2) $$\neg(\neg A \wedge B) \;\rightarrow\; (A \vee \neg B)$$

(hier allgemeiner B für A) ablehnen. Intuitionistisch weiterhin gültig ist dagegen

(3) $$\neg(\neg A \wedge B) \;\rightarrow\; (\neg\neg A \vee \neg B)$$

Damit nun aus (3) nicht (2) folgen soll, darf auch

DNB. $$\neg\neg A \rightarrow A$$

(Beseitigung doppelter Negationen) kein Theorem von **IL** sein.

Den Sinn der materialen Implikation in der klassischen Logik haben wir mit der Äquivalenz

(4) $$(A \rightarrow B) \;\leftrightarrow\; \neg A \vee B$$

wiedergegeben. Manchmal haben wir (4) auch als eine Definition des Pfeils $\rightarrow$ genommen. Auch (4) kann in **IL** nicht gelten. Denn aus $A \rightarrow A$ könnten wir dann TND gewinnen. Intuitionisten akzeptieren $A \rightarrow A$ und lehnen also die Äquivalenz (4) ab.

Diese Überlegungen müssen zunächst als eine Reihe von *ad hoc*-Manövern erscheinen, deren Ende nicht abzusehen ist. Intuitionisten bieten jedoch eine

Interpretation der Satzverknüpfungen an, die diese Manöver systematisch begründen; siehe z.B. Heyting [133, Kap. VII].

Vertreter der klassischen Logik werfen Intuitionisten vor, sie würden die Frage, ob eine Aussage wahr sei, verwechseln mit der Frage, ob wir über geeignete Mittel verfügen, die Wahrheit der Aussage zu erweisen. Intuitionisten antworten, daß hier gar keine Verwechslung vorliege. Die Zuschreibung von Wahrheit sei immer an die Verfügbarkeit solcher Mittel gebunden. Sie werfen ihrerseits den Klassikern vor, daß diese einen Wahrheitsbegriff verwenden, der über die menschliche Praxis der Vergabe des Wahrheitsprädikats offenbar hinausgeht und fordern den Klassiker heraus zu zeigen, welche Aspekte unserer Redepraxis zu der Annahme zwingen, daß wir Wahrheit in diesem stärkeren, klassischen Sinne verstehen. Das ist Dummetts *manifestation challenge*; siehe z.B. [65][66]. Den philosophischen Hintergrund dieser Herausforderung gibt offenbar eine Theorie der Bedeutung sprachlicher Ausdrücke ab, wonach die Frage nach der Bedeutung eines Ausdrucks erschöpfend durch die Beschreibung seines Gebrauchs in der Sprache beantwortet ist. Für die logischen Junktoren kann eine solche Bedeutungstheorie beispielsweise in der Angabe von Einführungs- und Beseitigungsregeln bestehen. Aus dieser Sicht liegt der Wert von Modellen, wie wir sie hier immer wieder betrachtet haben, nicht darin, daß sie eine *semantische* Theorie der betrachteten Sprache abgeben. Sie können allenfalls als Mittel taugen, um zum Beispiel etwas über den Begriff der logischen Wahrheit oder Folgerung herauszufinden. (Dies aber nur, solange wir uns dabei an die Bedingungen intuitionistischer Beweisführung halten.)

Intuitionistische Kripke-Modelle. Intuitionisten verknüpfen Wahrheit und Behauptbarkeit enger als dies Verfechter der klassischen Logik gemeinhin tun; manche Intuitionisten setzen die beiden Begriff in eins. Behauptbarkeit ist an die Verfügbarkeit von Evidenz gebunden. Es ist daher verständlich, warum unzureichende Evidenz für nicht-A kein zureichender Behauptungsgrund für A ist, d.h., warum DNB in diesem Sinne nicht gültig sein kann. Ebenso verfügen wir nicht in jeder Situation für jede Aussage über zureichende Gründe, die Aussage zu bejahen oder zu verneinen; manchmal läßt die verfügbare Evidenz beide Optionen offen und dann ist TND falsch.

Grzegorzek und Kripke haben unabhängig voneinander sehr ähnliche Vorschläge gemacht, wie man eine solche Motivation intuitionistischer Logik im Rahmen der klassischen Logik modellieren kann. Wir folgen hier dem Ansatz von Kripke [166]. Es sei der Vorsicht halber angemerkt, daß es eine "Standardinterpretation" der intuitionistischen Logik nicht gibt. Dazu gehen die Ansichten von Intuitionisten über eine richtige Beschreibung der

Absichten ihrer Unternehmung zu weit auseinander. Aber der jetzt zu beschreibende Ansatz ist abstrakt genug, um auf alle üblichen Kommentierungen der intutionistischen Logik zu passen. Er ist gewissermaßen ein Versuch, sich **IL** aus klassischer Perspektive verständlich zu machen.

Wir gehen aus von Kripke-Rahmen,

$$(W, R),$$

wobei W eine (nichtleere) Menge von Punkten ist, die wir jetzt "Erkenntnisstadien" oder "Situationen" nennen können. Die Relation $R \subseteq W \times W$ ordnet W im Sinne eines Evidenzzuwachses: Intuitiv soll Rab bedeuten, daß im Stadium b nicht mehr Evidenz als in a verfügbar ist. Dazu paßt, daß die Relation R *reflexiv* und *transitiv* ist, d.h. für alle $a, b, c \in W$:

$$Raa, \quad \text{und} \quad Rab \;\&\; Rbc \Rightarrow Rac.$$

Diese Bedingungen zusammengenommen beschreiben Erkenntnisfortschritt als einen sich möglicherweise verzweigenden Prozess.

Soweit haben wir nur reflexiv-transitive Kripke-Rahmen mit einer neuen Glossierung versehen. In der Sache macht das keinen Unterschied. Neues kommt jedoch hinzu, wenn wir diese Rahmen zur intuitionistischen Interpretation einer aussagenlogischen Sprache mit Junktoren $\{\neg, \wedge, \vee, \rightarrow\}$ heranziehen.

Als erstes stellen wir Interpretationen $I : \mathrm{ATM} \times W \longrightarrow \{0, 1\}$ unter die Bedingung:

$$\text{(Mon)} \qquad \text{wenn } I(P, a) = 1 \text{ und } Rab, \text{ dann } I(P, b) = 1.$$

Damit wird präzisiert, welche Art von Weiterentwicklung von Erkenntnisstadien die Relation R wiedergeben soll. Es geht um kumulative Fortentwicklungen; keine Information geht verloren. Die Bedingung (Mono) fordert das zunächst nur für atomare Information, aber wir werden gleich sehen, daß Monotonie ganz allgemein gilt.

Sodann interpretieren wir die Negation $\neg$ und die Implikation $\rightarrow$ neu.

$$(\neg_i) \qquad a \models \neg A \;\text{ gdw }\; \forall b: \text{ wenn } Rab, \text{ dann } b \not\models A.$$

Eine Aussage A an einem Punkt zu negieren, bedeutet auszuschließen, daß eine Zunahme von Evidenz A behauptbar machen kann – anders gesagt: daß am Punkt a genügend Evidenz verfügbar ist, A definitiv auszuschließen.

Da $\leq$ reflexiv ist, folgt aus $(\neg_i)$, daß jeder Punkt konsistent ist, d.h.

$$\text{wenn } a \models \neg A, \text{ dann } a \not\models A.$$

Erwartungsgemäß – anderenfalls TND! – gilt jedoch die Umkehrung, d.h. die Vollständigkeit der Punkte,

$$\text{Wenn } a \not\models A, \text{ dann } a \models \neg A,$$

nicht allgemein.

Auch die Implikation erfährt intuitionistisch eine neue Deutung:

$$(\to_i) \qquad a \models A \to B \;\text{ gdw }\; \forall b: \text{ wenn } Rab \text{ und } b \models A, \text{ dann } b \models B.$$

Eine intuitionistische Implikation gilt an einem Punkt, wenn sie nicht nur an diesem Punkt (Reflexivität!), sondern auch bei beliebiger Zunahme von Evidenz die Ablösung des Konsequens vom Antezedens erlaubt.

Die Bedingungen für die Konjunktion $\wedge$ und die Disjunktion $\vee$ sind wie in der klassischen Aussagenlogik. Wir müssen nun jedoch beide aufführen, da $\wedge$ und $\vee$ in **IL** nicht wechselseitig (d.h. unter Einsatz der Negation) definierbar sind.

$$(\wedge) \qquad a \models A \wedge B \text{ gdw } a \models A \text{ und } a \models B;$$

$$(\vee) \qquad a \models A \vee B \text{ gdw } a \models A \text{ oder } a \models B.$$

Anmerkung. Haben wir Negation und Implikation neu gedeutet oder haben wir neue Junktoren ($\neg_i$ und $\to_i$) unter alten Namen ($\neg$ und $\to$) eingeführt? Im ersten Fall, stünde die intuitionistische Interpretation in Konkurrenz zur klassischen, im zweiten Fall würden intuitionistische und klassische Logik über verschiedene Gegenstände sprechen – ganz im Sinne von Quines [236] Diktum: “Logikwechsel ist Themenwechsel”. In einem bestimmten Sinne sind beide Sichtweisen möglich. Einerseits können wir die Sache so sehen: Wir interpretieren Sprachen mit $\{\neg, \wedge, \vee, \to\}$, indem wir auf dieselbe Rahmenklasse jeweils *verschiedene* Erfüllungsrelationen aufsetzen. Die Mengen gültiger Formeln werden je nach Erfüllungsrelation verschieden ausfallen. Andererseits können wir aber auch eine Sprache mit $\{\neg, \neg_i, \wedge, \vee, \to, \to_i\}$ unter *einer* Erfüllungsrelation interpretieren. Wir erhalten so *eine* Menge gültiger Formeln und können dann verschiedene Fragmente dieser Menge betrachten; z.B. das klassische Fragment in $\{\neg, \wedge, \vee, \to\}$ oder das intuitionistische Fragment in $\{\neg_i, \wedge, \vee, \to_i\}$.[36]

[36] Unsere Notation kann bis auf weiteres unentschieden zwischen den beiden Sichtweisen bleiben. Streichen wir in “($\neg_i$)” und “($\to_i$)” die Indizes, so mag das auf eine Neuinterpretation hindeuten; versehen wir in diesen Klauseln die Junktoren mit Indizes i, dann kann das als Einführung neuer Junktoren verstanden werden. Wenn wir weiter unten (pp. 187ff.) zur Frage der Übersetzbarkeit zwischen klassischer und intuitionistischer Logik kommen, ist es manchmal klarer oder gar geboten, die zweite Sichtweise einzunehmen.

Daß es in der Kontroverse zwischen Klassikern und Intuitionisten im Kern um die Deutung der Negation geht, kann, nach dem bisher Gesagten, nicht ganz richtig sein. Zwar ist es so, daß wenn wir Heytings Axiomatisierung von **IL** um das Schema DNB ergänzen, wir eine Axiomatisierung von **KL** erhalten. Das Prinzip der doppelten Negationsbeseitigung separiert in diesem Sinne die beiden Logiken. Aber genau in diesem Sinne tut das auch das rein implikative Peirce-Schema $((A \to B) \to A) \to A$. Mindestens ebenso wichtig ist also das richtige Verständnis der Implikation. Tatsächlich könnten wir in der Beschreibung intuitionistischer Kripke-Modelle ganz ohne eine Interpretationsbedingung für die Negation auskommen. Stattdessen setzen wir die *Falsum*-Konstante $\bot$ ein mit der Bedingung – die in der klassischen Logik nicht anders ist –

$$(\bot) \qquad \forall a: \ a \not\models \bot,$$

d.h. an keinem Punkt ist $\bot$ behauptbar. (Die Widerspruchskonstante $\bot$ kann in diesem Fall natürlich nicht mit $P \wedge \neg P$ definiert sein, für irgendein Atom P. Die Konstante $\bot$ muß vielmehr als primitiv aufgefaßt werden; in bestimmten Fällen sind Definitionen wie $\bot = (0 = 1)$ möglich.) Wenn wir jetzt die Negation so definieren, wie wir das auch in der klassischen Logik tun können, nämlich

$$(\text{Def. } \neg) \qquad \neg A \ = \ A \to \bot,$$

dann ist es nun allein die Interpretation der Implikation, die darüber entscheidet, ob die Negation die klassischen oder intuitionistischen Eigenschaften hat. Unschwer ist zu sehen, daß (Def. $\neg$) im Verein mit der intuitionistischen Bedingung $(\to_i)$ die intuitionistische Bedingung $(\neg_i)$ ergibt.

Wir schließen die Beschreibung der Klasse intuitionistischer Modelle mit der Definition von Wahrheit und Gültigkeit. Wir wollen hier mit $\mathbb{I}$ die Klasse der reflexiven und transitiven Rahmen bezeichnen; $Mod_i(\mathbb{I})$, mit den Definitionen $(\neg_i)$ und $(\to_i)$, ist dann die Klasse intuitionistischer Modelle auf $\mathbb{I}$.

- Eine Formel ist genau dann *wahr in einem intuitionistischen Modell*, wenn die Formel an allen Punkten des Modells wahr ist.
- Eine Formel ist genau dann *intuitionistisch gültig*, wenn sie in allen Modellen in $Mod_i(\mathbb{I})$ wahr ist. Mit $Th_i(\mathbb{I})$ bezeichnen wir die intuitionistische Theorie der Rahmenklasse $\mathbb{I}$, d.h. die Menge der intuitionistisch gültigen Formeln.

* * *

Intuitionistische Aussagenlogik **IL** (Heyting)

I1.	$A \to A \wedge A$
I2.	$A \wedge B \to B \wedge A$
I3.	$(A \to B) \to (A \wedge C \to B \wedge C)$
I4.	$(A \to B) \wedge (B \to C) \to (A \to C)$
I5.	$A \to (B \to A)$
I6.	$A \wedge (A \to B) \to B$
I7.	$A \to A \vee B$
I8.	$A \vee B \to B \vee A$
I9.	$(A \to C) \wedge (B \to C) \to (A \vee B \to C)$
I10.	$\neg A \to (A \to B)$
I11.	$(A \to B) \wedge (A \to \neg B) \to \neg A$
MP.	$\dfrac{A \quad A \to B}{B}$

In Heyting [133, p. 101] finden wir eine Axiomatisierung der intuitionistischen Aussagenlogik **IL**. Fortan soll mit **IL** diese axiomatische Repräsentation der intuitionistischen Aussagenlogik gemeint sein. Wir wollen den folgenden Satz beweisen.

Satz 77. *Eine Formel A ist genau dann ein Theorem von* **IL**, *wenn* $A \in Th_i(\mathbb{I})$.

Zunächst beobachten wir jedoch, daß sich die Monotoniebedingung (Mon) für Atome über den gesamten Bereich der Formeln fortsetzt:

Lemma 78. (Monotonie) *In jedem intuitionistischen Modell gilt für alle Formeln A: Wenn $a \models A$ und Rab, dann $b \models A$.*

Beweis. Einfache Induktion über den Aufbau von A. Der atomare Fall ist durch die Monotonie-Bedingung für atomare Formeln gegeben. Für Konjunktionen und Disjunktionen folgt das Resultat unmittelbar aus der Induktionsannahme. Für Negationen und Implikationen machen wir ohne die Induktionsannahme Gebrauch von der Transitivität der Relation R. ∎

Den Beweis des Satzes 77 von links nach rechts, d.h. in der Richtigkeitsrichtung, führt man über eine Routine-Verifikation der Axiome und Regeln von **IL** in $Mod_i(\mathbb{I})$.

Für die Vollständigkeitsrichtung bedienen wir uns wieder kanonischer Modelle. Im Falle modaler Erweiterungen der klassischen Logik basieren diese auf Formelmengen, die maximal und konsistent im Sinne der betrachteten Logik sind. Das garantiert, daß diese Formelmenge sich wie Punkte in einem Modell verhalten. Die Punkte in einem intuitionistischen Modell sind jedoch nicht vollständig: Aus $a \not\models A$ folgt nicht, daß $a \models \neg A$. Da Maximalität (zusammen mit Konsistenz) Vollständigkeit impliziert, so sollte das kanonischen Modell für **IL** nicht aus maximalen Formelmengen bestehen. Für den Vollständigkeitsbeweis dienen vielmehr solche Formelmengen dem Zweck, die *prim*, *konsistent* und *deduktiv abgeschlossen* ("pkd") im Sinne von **IL** sind.

DEFINITION 79. Eine Formelmenge X ist
prim gdw wenn $A \vee B \in X$, dann $A \in X$ oder $B \in X$;
konsistent gdw $\perp \notin X$;
deduktiv abgeschlossen (i.S.v. **IL**) gdw
wenn $A_1, ..., A_n \in X$ und $A_1, ..., A_n \vdash A \in$ **IL**, dann $A \in X$.
X ist genau dann *pkd*, wenn X alle drei Eigenschaften besitzt.

Im Anschluß an den Beweis von Satz 77 werden wir zeigen, daß **IL**, im Gegensatz etwa zur klassischen Logik, selbst prim ist – die *Disjunktionseigenschaft* hat, wie man auch sagt.

Das folgende Lemma ist eine Adaptation des Lindenbaum-Lemmas.

LEMMA 80. *Jede konsistente Menge läßt sich zu einer pkd'en Menge erweitern.*

BEWEIS. Wie im Beweis des Lindenbaum-Lemmas (siehe das Einleitungskapitel) zählen wir alle Formeln ab (beginnend mit 1) und beginnen die Konstruktion einer pkd'en Erweiterung einer konsistenten Menge X mit $X_0 = X$. Im weiteren verfahren wir nun so:

$$X_{n+1} = \begin{cases} X_n, \text{ falls } X_n, A_n \vdash \perp; \text{ anderenfalls:} \\ \begin{cases} X_n \cup \{A_n, B\}, \text{ falls } A_n = B \vee C, \\ X_n \cup \{A_n\} \text{ anderenfalls.} \end{cases} \end{cases}$$

Schließlich sei X^* die Vereinigung aller $X_{n\leq 0}$. In der zweiten Zeile der Definition von $X_n + 1$ stellen wir sicher, daß keine Disjunktion ohne eines ihrer Disjunkte in die Konstruktion eintritt. Daß dadurch nicht die Eigenschaft der Konsistenz gefährdet wird, die wir in der ersten Zeile gewährleisten wollen, zeigt das folgende Argument. Unter der Annahme

(1) $X, B \vee C \nvdash \perp$

nehmen wir ferner – für eine *reductio* – an, daß

(2) $X, B \vee C, B \vdash \bot$.

Da sowohl $B \vdash B \vee C$ also auch $C \vdash B \vee C$, so erhalten wir aus (2) durch Schnitt

(3) $X, B \vdash \bot$ und $X, C \vdash \bot$.

Aus (3) folgt wiederum

(4) $X, B \vee C \vdash \bot$,

was unserer Annahme (1) widerspricht. Es läßt sich leicht zeigen, daß X^* deduktiv abgeschlossen, konsistent und prim ist. ∎

Pkd'e Mengen haben Eigenschaften, die auf die intuitionistischen Interpretationsbedingungen genau passen, wie das folgende Lemma zeigt.

LEMMA 81. *Es sei X eine pkd'e Menge. Dann gilt für alle Formeln A und B:*

1. *$A \wedge B \in X$ gdw $A \in X$ und $B \in X$;*
2. *$A \vee B \in X$ gdw $A \in X$ oder $B \in X$;*
3. *$\neg A \in X$ gdw $A \notin Y$, für alle pkd'en Erweiterungen Y von X;*
4. *$A \rightarrow B \in X$ gdw für alle pkd'en Erweiterungen Y von X: wenn $A \in Y$ dann $B \in Y$.*

BEWEIS. *Übung.* ∎

DEFINITION 82. Das kanonische Modell für **IL**,

$$(W^{\mathbf{IL}}, R^{\mathbf{IL}}, I^{\mathbf{IL}}),$$

besteht aus

$W^{\mathbf{IL}}$: der Familie aller pkd'en Mengen,
$R^{\mathbf{IL}} = \subseteq$,
$I^{\mathbf{IL}} : \mathrm{ATM} \longrightarrow \{0,1\}$ so, daß $(\forall a \in W^{\mathbf{IL}})\ I^{\mathbf{IL}}(P, a) = 1$ gdw $P \in a$.

(Im folgenden lassen wir die Superskripte wieder weg.) Aus der Definition folgt recht unmittelbar, daß das kanonische Modell zur Klasse $Mod_i(\mathbb{I})$ gehört. Unter Rückgriff auf Lemma 81 läßt sich nun leicht das folgende Analogon zu Lemma 12 (dem "Hauptsatz") beweisen:

LEMMA 83. *Es sei (W, R, I) das kanonische Modell für* **IL**. *Es sei eine Relation $\models \subseteq S \times FML$ so definiert:*

$$a \models A \text{ gdw } A \in a.$$

Dann ist die Relation $\models$ eine Erfüllungsrelation im intuitionistischen Sinne.

Um schließlich den Beweis von Satz 77 zu vervollständigen, nehmen wir an, daß $A \notin$ **IL**. Dann können wir ausgehend von **IL** nach der adaptierten

Lindenbaum-Methode eine pkd'e Erweiterung X konstruieren mit $A \notin X$. X ist somit ein Punkt im kanonischen Modell für **IL**, welcher A falsifiziert. Also ist A nicht in $Th_i(\mathbb{I})$, d.h. **IL** ist vollständig bezüglich der Klasse intuitionistischer Modelle.

Für den Intuitionisten ist eine Disjunktion $A \vee B$ im Rahmen einer Theorie nur dann behauptbar, wenn eines der Disjunkte, A oder B, in dieser Theorie behauptbar ist. Das sollte insbesondere auch der Fall sein, wenn es sich um die Theorie logischer Wahrheit handelt. Es wäre merkwürdig, wenn Intuitionisten z.B. das *Tertium Non Datur* auf der Grundlage dieses Verständnisses der Disjunktion zurückweisen wollten, für die intuitionistische Aussagenlogik aber disjunktive Theoreme zulassen müßten, ohne das eines der Disjunkte ein Theorem wäre. Auf der Grundlage von Satz 77 können wir jetzt zeigen, daß **IL** tatsächlich die gewünschte Disjunktionseigenschaft hat.

Satz 84. **IL** *ist prim:* $A \vee B \in \mathbf{IL}$ *gdw* $A \in \mathbf{IL}$ *oder* $B \in \mathbf{IL}$.

Beweis. (LR): Angenommen $A \notin \mathbf{IL}$ und $B \notin \mathbf{IL}$. (Wir zeigen, daß es ein Modell $\mathcal{M}$ mit einem Punkt c gibt so, daß $c \not\models A \vee B$. Nach Satz 77 (LR) folgt dann, daß $A \vee B \notin \mathbf{IL}$.) Unter der Annahme gibt es nach Satz 77 (RL) Modelle $\mathcal{M}_1 = (W_1, R_1, I_1)$ und $\mathcal{M}_1 = (W_1, R_1, I_1)$ mit jeweils Punkten a und b so, daß $\mathcal{M}_1, a \not\models A$ und $\mathcal{M}_2, b \not\models B$. Nun erzeugen wir zunächst aus a bzw. b die Teilmodelle $\mathcal{M}_1^a$ und $\mathcal{M}_2^b$ für die natürlich ebenfalls gilt

$$(*) \qquad \mathcal{M}_1^a, a \not\models A \text{ und } \mathcal{M}_2^b, b \not\models B.$$

Sodann fügen wir $\mathcal{M}_1^a$ und $\mathcal{M}_2^b$ zusammen und fügen einen weiteren Punkt c so hinzu:

$$\begin{aligned}
\mathcal{M} &= (W, R, I) \text{ mit} \\
W &= W_1^a \cup W_2^b \cup \{c\}, \\
R &= R_1^a \cup R_2^b \cup \{(c, x) : x \in W\}, \\
I &= I_1^a \cup I_2^b.
\end{aligned}$$

(Während $\mathcal{M}_1$ und $\mathcal{M}_2$ Modelle mit Wurzel a bzw. b sind, ist $\mathcal{M}$ nun ein Modell mit Wurzel c und Rca sowie Rcb.) Da M_1^a und M_2^b aus $\mathcal{M}$ erzeugbare Teilmodelle sind, so folgt aus $(*)$, daß $\mathcal{M}, a \not\models A$ und $\mathcal{M}, b \not\models B$. Aber dann (Monotonie!) auch $\mathcal{M}, c \not\models A$ und $\mathcal{M}, c \not\models B$, d.h. $\mathcal{M}, c \not\models A \vee B$. ∎

Johanssons Minimalkalkül, JL. Heyting [133, p. 102] kommentiert das Axiom

$$\text{(I10.)} \qquad \neg A \to (A \to B)$$

so:

> Axiom [I10] may not seem intuitively clear. As a matter of fact, it adds to the precision of the definition of implication. You remember that $A \to B$ can be asserted if and only if we possess a construction which, joined to the construction of A, would prove B. Now suppose that $\vdash \neg A$, that is, we have deduced a contradiction from the supposition that A were carried out. Then, in a sense, this can be considered as a construction, which joined to a proof of A (which cannot exist) leads to a proof of B. I shall interpret the implication in this wider sense.
>
> A system of intuitionistic logic in which $\to$ is interpreted in the narrower sense and in which, accordingly, [I10] is rejected as an axiom, has been developed by Johansson in his "minimal calculus".

In einem Sinne also, ist das Minimalkalkül **JL** (siehe [157]) die strengere Umsetzung der intuitionistischen Kritik an der klassischen Auffassung der Aussagenlogik.

Die Logik **JL** hat ferner den Vorzug, daß sie inkonsistente Theorien nicht völlig trivialisiert. Wenn es in einer Theorie eine Instanz für sowohl A als auch für $\neg A$ gibt, dann läßt sich nach Schema I10 (und zweimaligem Modus Ponens) daraus jede beliebige Formel B ableiten. Das Minimalkalkül **JL** trivialisiert inkonsistente Theorien nicht. Eine Logik, die zwischen inkonsistenten und trivialen Theorien unterscheiden kann, nennt man *parakonsistent*. (Solche Logiken werden ausführlich in Kapitel V behandelt.) Das System **JL** ist also ein Beispiel einer parakonsistenten Logik – jedoch nur dem Buchstaben nach. Denn obwohl das Schema I10 kein Theorem von **JL** ist, können wir in **JL** das folgende Schema ableiten:

JOH. $$\neg A \to (A \to \neg B).$$

Das bedeutet aber, daß unter minimaler Konsequenz abgeschlossene inkonsistente Theorien "halbtrivial" sind: Sie enthalten von allen Formeln der Sprache deren Negationen, d.h. die "negative Hälfte" der Formelmenge. Dennoch: der positive Teil der Konsequenzenmenge einer inkonsistenten Theorie ist nicht trivial. Dieser enthält Informationen, die für manche Zwecke nützlich sein können.

Wir haben auf p. 176 beobachtet, daß alle Punkte in intuitionistischen Modellen konsistent sind. Das schließt aus, daß in einem Modell Formeln A und $\neg A$ wahr werden. Das Schema $\neg A \to (A \to B)$ muß also unter der Bedingung, daß es keine inkonsistenten Punkte gibt, auf triviale Weise gültig sein. Wenn wir die Möglichkeit solcher Punkte jedoch einräumen, dann können wir das Schema falsifizieren. Das ist die Grundidee der folgenden Modellierung des Minimalkalküls.

Wir erweitern die bisher betrachteten Strukturen (W, R) um eine Funktion *, welche W in eine Teilmenge von W abbildet. Die Rahmen sehen jetzt so aus:

$$(W, {}^*, R).$$

Die Idee ist es, daß die Funktion * aus W eine Menge W^* "normaler" Punkte aussondert. Was das bedeutet, wird durch die Erfüllungsbedingung für negierte Formeln bestimmt: Für alle $a \in W$,

$$(\neg_j) \qquad a \models \neg A \text{ gdw } \forall b : b \in W^* \ \& \ Rab \ \Rightarrow \ b \not\models A.$$

Alle anderen Bedingungen, einschließlich der Monotonie sowie der Reflexivität und Transitivität, bleiben wie für intuitionistische Rahmen bzw. Modelle. Insbesondere können wir auch für diese Modelle zeigen, daß Monotonie nicht nur Atome, sondern alle Formeln regiert (siehe das Monotonie-Lemma 78). Aus Monotonie und der Reflexivität von R folgt unmittelbar für alle Formeln A,

$$\text{wenn } a \in W^* \text{ und } a \models \neg A, \text{ dann } a \not\models A.$$

Normalität impliziert also Konsistenz. Wenn a jedoch nicht in S^* ist, dann verbietet $(\neg_j)$ nicht, daß wir sowohl $a \models \neg A$ als auch $a \models A$ haben. Genau an dieser Möglichkeit scheitert die Verifikation von I10 wie folgt:

(1)	Rab	Annahme
(2)	$b \models \neg A$	Ann. (zz $b \models A \to B$)
(3)	Rbc	Ann.
(4)	$c \models A$	Ann. (zz $c \models B$)
(5)	$c \models \neg A$	2,3, Monotonie
(6)	$\forall x : Rcx \ \& \ x \in W^* \ \Rightarrow \ x \not\models A$	5

Wenn nun $x = c$ und $c \notin W^*$ mit $c \models B$, dann haben wir ein Muster für ein Gegenbeispiel zu $a \models \neg A \to (A \to B)$. Nehmen wir aber an, daß statt $c \models B$ zu zeigen ist, daß $c \models \neg B$, d.h.

$$\forall x : Rcx \ \& \ x \in W^* \ \Rightarrow \ x \not\models B,$$

dann können wir so fortsetzen und die Gültigkeit von JOH zeigen:

(7)	Rcd	Ann.
(6)	$d \in W^*$	Ann. (zz $d \not\models B$)
(7)	$d \models A$	4,7, Monotonie
(8)	$d \not\models A$	6,7

Damit erreichen wir einen Widerspruch aus dem wir trivialerweise auf $d \not\models B$ schließen dürfen.

Der Nachweis, daß **JL** richtig und vollständig im Hinblick auf die gerade beschriebene Klasse von minimallogischen Modellen ist, verläuft im wesentlichen wie für **IL** (Übung!).

Subklassische Erweiterungen von IL. Der Minimalkalkül **JL** ist eine naheliegende Abschwächung von **IL**, eine subintuitionistische Logik. In mancherlei Hinsicht liegt er aber ebenso nahe, die intuitionistische Logik etwas anzureichern, ohne dabei gleich wieder bei der klassischen Logik anzukommen. Diesen Bereich zwischen intuitionistischer und klassischer Logik nennt man den Bereich der *intermediären Logiken.* Wir erwähnen hier zwei prominente Schemata, welche zu subklassischen Erweiterungen von **IL** führen. Diese ergeben sich recht natürlich aus der Betrachtung der Modelle. Beide Erweiterungen bewahren übrigens die Disjunktionseigenschaft der intuitionistischen Logik, d.h. die jeweiligen Theoremmengen sind prim.

Die *Dummett-Formel*

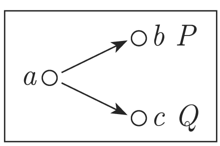
DUM. $$(A \to B) \vee (B \to A)$$

korrespondiert mit der Bedingung der Rechts-Linearität,

(R-Lin) wenn Rab und Rac, dann Rbc oder Rcb.

Die Bedingung fordert, daß alle Punkte sich in Ketten einfügen, schließt aber nicht aus, daß es mehrere solcher Ketten gibt; vgl. die Diskussion von (R-Lin) im Rahmen der Zeitlogik. Um die Korrespondenz zu beweisen, zeigen wir, 1. daß DUM nicht gilt, wenn ein Rahmen die Bedingung nicht erfüllt und, umgekehrt, 2. daß die Bedingung hinreichend für die Gültigkeit von DUM ist.

Ad 1. Es seien in einem Rahmen die Punkte a, b, und c derart, daß Rab und Rac, jedoch mögen b und c unverbunden sein durch die Relation R. (Alle Punkt sind reflexiv.) Wir betrachten eine Bewertung I mit $I(P,b) = 1 = I(Q,c)$ und $I(Q,b) = 0 = I(P,c)$ – in einem Bild:



Dann ist $a \not\models P \to Q$ und $a \not\models Q \to P$, also $a \not\models$ DUM.

Ad 2.

(1)	$a \not\models A \to B$	Ann.
(2)	$a \not\models B \to A$	Ann.
(3)	$\exists b : Rab$ und $b \models A$ und $b \not\models B$	1
(4)	$\exists c : Rac$ und $c \models B$ und $c \not\models A$	2
(5)	Rbc oder Rcb	3,4, (R-Lin)
(6)	Rbc	5: erster Fall
(7)	$b \not\models A$	4,6, Monotonie
(8)	Widerspruch	erster Fall, aus 3,7
(9)	Rcb	5: zweiter Fall
(10)	$c \not\models B$	3,9, Monotonie
(11)	Widerspruch	zweiter Fall, aus 4,10
(12)	Widerspruch	aus den Ann. 1,2

Das *Schwache Tertium Non Datur*

STND. $\neg A \vee \neg\neg A$

korrespondiert mit der Bedingung der Konvergenz,

(Konv) wenn Rab und Rac, dann $\exists d : Rbd$ und Rcd.

Daß Konvergenz notwendig ist für die Gültigkeit von STND erkennen wir, wenn wir einen Rahmen betrachten, in dem a, b, c nicht in einem Punkt konvergieren. Dann kann $\neg P \vee \neg\neg P$ so am Punkte a falsch werden: $I(P, b) = 1$ und $b \not\models \neg P$. Letzteres ist der Fall, wenn P an einer Fortsetzung von b wahr ist (zum Beispiel am Punkt b selbst – Reflexivität!), ohne daß es nun Fortsetzungen x von b oder c gibt so, daß die eine die andere sehen kann. Das Diagramm für DUM illustriert die Situation, wenn wir uns die Beschriftung Q am Punkt c wegdenken. Es läßt sich leicht zeigen, daß die Existenz konvergierender Fortsetzungen x hinreichend für STND ist.

Wie stellen wir fest, daß eine vorgeschlagene Erweiterung von **IL** in den Bereich der intermediären Logiken fällt und nicht etwa die klassische Logik **KL** produziert? Wir bräuchten ein "Sieb", das diejenigen Tautologien aussondert, welche einerseits unabhängig von **IL** sind und andererseits **IL** nicht nach **KL** umschlagen läßt. Dazu bedienen wir uns Matrizen, wie wir sie schon benutzt haben, um die Unabhängigkeit von Schemata im Hinblick auf ein axiomatisches System zu beweisen – nur, daß wir diese Matrizen nun auf die gerade erwähnte stärkere Eigenschaft testen müssen. Die folgenden Matrizen gehen auf Jankov [154] zurück:

$\neg$	
1	3
2	3
3	1

$\wedge$	1	2	3
1	1	2	3
2	2	2	3
3	3	3	3

$\vee$	1	2	3
1	1	1	1
2	1	2	2
3	1	2	3

$\to$	1	2	3
1	1	2	3
2	1	1	3
3	1	1	1

Der designierte Wert ist 1. Eine Logik $\mathbf{IL} + A$ ist genau dann intermediär (d.h. $\mathbf{IL} + A \neq \mathbf{KL}$) wenn A eine Tautologie ist, die in den Matrizen gültig ist, d.h. bei jeder Bewertung der Teilformeln den designierten Wert 1 erhält. (Daß diese Matrizen dem Zweck dienen, möge der Leser an DUM und STND überprüfen.)

Übersetzung in S4. Es gibt keine Wahrheitsfunktionen, welche die intuitionistischen Junktoren $\neg_i$ und $\to_i$ darstellen könnten. So gibt es auch keine Übersetzung * von einer intuitionistischen Sprache $\mathcal{L}_i$ mit $\{\neg_i, \wedge, \vee, \to_i\}$ in eine klassische Sprache mit $\{\neg, \wedge, \vee, \to\}$ so, daß für jede Formel A der intuitionistischen Sprache gilt, daß $A \in \mathbf{IL}$ genau dann, wenn $A^* \in \mathbf{KL}$.

Anders sieht die Sache jedoch aus, wenn wir in die modale Erweiterung $\mathcal{L}^\Box$ einer klassischen Sprache übersetzen. In einer Notiz aus dem Jahre 1933 [112] bemerkt Gödel:

> Man kann den Heytingschen Aussagenkalkül mittels der Begriffe des gewöhnlichen Aussagenkalküls und des Begriffes "A ist beweisbar" (bezeichnet mit $\Box A$) interpretieren, wenn man für den letzteren das folgende Axiomensystem [...] annimmt: [Es folgt eine Axiomatisierung von **S4**].

Darauf gibt Gödel eine Interpretation (Übersetzung) an, beobachtet, daß die Übersetzung jedes Theorems von **IL** ein Theorem von **S4** ist, und vermutet, daß das auch umgekehrt gilt. [37] Mehrere Übersetzungen sind möglich. Aus der Darstellung intuitionistischer Kripke-Modelle ergibt sich auf natürliche Weise die folgende Übersetzung t von $\mathcal{L}_i$ nach $\mathcal{L}^\Box$ (erstmals in McKinsey und Tarski [211]):

Def. t

$$\begin{aligned} P^t &= \Box P; \\ (\neg_i A)^t &= \Box\neg A^t; \\ (A \wedge B)^t &= A^t \wedge B^t; \\ (A \vee B)^t &= A^t \vee B^t; \\ (A \to_i B)^t &= \Box(A^t \to B^t). \end{aligned}$$

[37] Chagrov und Zakharyaschew [50, p. 57] berichten, daß der russische Logiker Orlov schon 1928 eine Übersetzung von **IL** nach **S4** gefunden hat.

Satz 85. *Eine Formel A aus $\mathcal{L}_i$ ist genau dann ein Theorem von* **IL**, *wenn A^t ein Theorem von* **S4** *ist.*

Beweis. Es sei $\mathcal{M} = (\mathcal{R}, I)$ ein modales Modell auf einem reflexiv-transitivem Rahmen $\mathcal{R}$. Wir definieren ein Modell $\mathcal{M}_i = (\mathcal{R}, I_i)$ mit

$$I_i(P,a) = 1 \text{ gdw } \mathcal{M}, a \models \Box P.$$

Es ist klar, daß $\mathcal{M}_i$ die Monotoniebedingung für atomare Formeln erfüllt. Es sei nun $\models_m$ die Erweiterung von I für alle Formeln in $\mathcal{L}^\Box$ und $\models_i$ sei die Erweiterung von I_i für alle Formeln in $\mathcal{L}_i$. Insbesondere gilt so nach der Definition von I_i, daß

$$a \models_i P \text{ gdw } a \models_m \Box P.$$

Beobachtung: Für jede Formel in $\mathcal{L}_i$ und alle Punkte a:

$$a \models_i A \text{ gdw } a \models_m A^t.$$

Der Beweis verfährt per Induktion über A. Die Basis – A ist ein Atom – ist durch die Definition von I_i gegeben. — *Ad* $A = \neg_i B$:

$$\begin{aligned} a \models_i \neg_i B &\text{ gdw } \forall b : Rab \Rightarrow b \not\models_i B && (\neg_i)\\ &\text{ gdw } \forall b : Rab \Rightarrow b \not\models_m B^t && \text{IA}\\ &\text{ gdw } \forall b : Rab \Rightarrow b \models_m \neg B^t && (\neg)\\ &\text{ gdw } a \models \Box\neg B^t && (\Box) \end{aligned}$$

Die Fälle für $\rightarrow_i$, $\wedge$ und $\vee$ sind ähnlich einfach.

Die Beobachtung gilt für beliebige Modellpaare $(\mathcal{R}, I)$ und $(\mathcal{R}, I_i)$ auf beliebigen reflexiv-transitiven (S4-) Rahmen $\mathbb{R}$. Wenn nun $A \notin$ **IL**, dann gibt es einen Punkt a in einem Modell auf einem S4-Rahmen so, daß $a \not\models_i A$. Daraus folgt, daß $a \not\models_m A^t$, d.h. $A \notin$ **S4**. Sei umgekehrt $A^t \notin$ **S4**, dann gibt es einen Punkt a in einem Modell $\mathcal{M}$ auf einem S4-Rahmen mit $a \not\models_m A$. Daraus folgt, daß in dem intuitionistischen Modell $\mathcal{M}_i, a \not\models_i A$, d.h. $A \notin$ **IL**.

■

Glivenkos Theorem. Die Einbettung von **IL** in **S4** erlaubt es, **IL** aus klassischer Perspektive zu verstehen. Intuitionistisch gültige Formeln erweisen sich – d.h. übersetzt – als gültige Formeln in einer modalen Erweiterung der klassischen Logik. Kann auch der Intuitionist auf ähnliche Weise für sich übersetzen, was der Klassiker unter einer logischen Wahrheit versteht? Ja, sehr einfach: Sagt der Klassiker, A sei eine logische Wahrheit, so versteht der Intuitionist, $\neg\neg A$ sei logisch wahr. Das ist der Inhalt von Glivenkos Theorem. Nach einem "traditionellen" Beweis aus der Axiomatik legen wir einen weiteren Beweis vor, der von einigen modelltheoretischen Mitteln Gebrauch macht, die wir in diesem Kapitel kennengelernt haben.

SATZ 86. (Glivenko 1927) $A \in \mathbf{KL}$ *gdw* $\neg\neg A \in \mathbf{IL}$.

ERSTER BEWEIS. Aus $\mathbf{IL} \subseteq \mathbf{KL}$ und der klassischen DNB folgt sofort die Implikation von rechts nach links. Wir wollen nun annehmen, **IL** sei durch (I1–11) und MP axiomatisiert. Wir wissen, daß wir eine Axiomatisierung von **KL** erhalten, wenn wir zu diesen Axiomen das Schema DNB, $\neg\neg A \to A$, hinzunehmen. Wir setzen hier zwei Theoreme von IL als gegeben voraus:

DNE. $$A \to \neg\neg A$$

DNB′ $$\neg\neg(\neg\neg A \to A)$$

Ferner ist $\neg(\neg A \to A) \to (\neg\neg A \to \neg A) \in \mathbf{IL}$, weshalb insbesondere der Schluß

MP′. $$\frac{\neg\neg(\neg\neg A \to A) \quad \neg\neg\neg\neg A}{\neg\neg A}$$

in **IL** zulässig ist. (Der Leser möge sich an Ableitungen versuchen oder nach Satz 77 verfahren.)

Angenommen nun, $A \in \mathbf{KL}$. Dann gibt es eine Ableitung (in der gerade genannten Axiomatisierung von **KL**)

$$\Delta = (A_0, \ldots, A_n) \text{ mit } A_n = A.$$

Wenn in Δ keine Instanz von DNB vorkommt, dann ist Δ eine Ableitung in **IL** von A und also – aufgrund von DNE und MP – auch eine von $\neg\neg A$, *qed*.

Wenn DNB in Δ vorkommt, dann betrachten wir

$$\Delta' = (\neg\neg A_0, \ldots, \neg\neg A_n).$$

Vorkommen von $\neg\neg B \to B$ in Δ sind in Δ' ersetzt durch Instanzen von DNB′, welche als Prämissen für MP′ dienen. Also ist Δ' eine Ableitung in **IL** von $\neg\neg A$, *qed*. ∎

ZWEITER BEWEIS. Zunächst wollen wir festhalten, daß **IL**, genau wie **S4**, die endliche Rahmen- und Modelleigenschaft hat, d.h. (vgl. Definition 65) wenn eine Formel kein Theorem von **IL** ist, dann gibt es ein endliches Kripke-Modell für **IL**, welches die Formel widerlegt.

Angenommen nun, $A \in \mathbf{KL}$ und – *reductio*! – $\neg\neg A \notin \mathbf{IL}$. Dann gibt es ein endliches Modell $\mathcal{M}$ mit einem Punkt a so, daß $a \not\models \neg\neg A$; d.h. $\exists b$ so, daß Rab und $b \models \neg A$. Es sei nun c ein terminaler Nachfolger von b unter R (ein Punkt ohne R-Nachfolger im transitiven Abschluß von $R(b)$). Nach dem Monotonie-Lemma folgt aus $b \models \neg A$, daß $c \models \neg A$ und also ($Rcc!$) $(*)$ $c \not\models A$ in $\mathcal{M}$. Nun betrachten wir das aus c erzeugte Teilmodell $\mathcal{M}^c$. Nach dem Satz 32 über erzeugte Teilmodelle gilt $(*)$ auch für $\mathcal{M}^c$. Da c terminal ist, enthält $\mathcal{M}^c$ nur den einen Punkt c. Gültigkeit in einpunktigen Kripke-Modellen koinzidiert aber mit klassischer Gültigkeit. Aus der Annahme $A \in \mathbf{KL}$ folgt so, daß $c \models A$ in $\mathcal{M}^c$ – im Widerspruch zu $(*)$. ∎

Glivenkos Satz ist der Ausgangspunkt für eine Reihe von Beobachtungen über das Verhältnis von klassischer und intuitionistischer Logik.

KOROLLAR 87. $\neg A \in \mathbf{KL}$ *gdw* $\neg A \in \mathbf{IL}$

BEWEIS. (LR): Wenn $\neg A \in \mathbf{KL}$, dann (nach Glivenkos Satz) $\neg\neg\neg A \in \mathbf{IL}$. Aus der Beobachtung, daß das Schema

TNB. $$\neg\neg\neg A \to \neg A$$

ein Theorem von $\mathbf{IL}$ ist, folgt, daß $\neg A \in \mathbf{IL}$. ∎

KOROLLAR 88. *Für jede Formel, in der nur $\neg$ und $\wedge$ vorkommt: $A \in \mathbf{KL}$ gdw $A \in \mathbf{IL}$.*

BEWEIS. (LR): Jede Formel A in $\{\neg, \wedge\}$ ist von der Form

$$A_0 \wedge \cdots \wedge A_n \ (0 \leq n),$$

wobei jedes der A_i $(0 \leq i \leq n)$ entweder ein Atom oder eine Negation $\neg B$ ist. Angenommen nun, $A \in \mathbf{KL}$. Dann sind alle $A_i \in \mathbf{KL}$, wobei A_i kein Atom sein kann. Also ist jedes A_i von der Form $\neg B$. Es folgt aus dem vorigen Korollar, daß jedes $A_i \in \mathbf{IL}$. Da die Menge der Theoreme von $\mathbf{IL}$ unter Konjunktion abgeschlossen ist, so ist auch $A \in \mathbf{IL}$. ∎

KOROLLAR 89. *$A \in \mathbf{KL}$ gdw $A^s \in \mathbf{IL}$, wobei $s : \mathcal{L} \longrightarrow \mathcal{L}_i$ die folgende Übersetzung ist:*

Def. s
$$\begin{aligned} P^s &= P \\ (\neg A)^s &= \neg A^s \\ (A \wedge B)^s &= A^s \wedge B^s \\ (A \vee B)^s &= \neg(\neg A^s \wedge \neg B^s) \\ (A \to B)^s &= \neg(A^s \wedge \neg B^s) \end{aligned}$$

BEWEIS. Die Übersetzung überführt Formeln in $\mathcal{L}$ in das $\{\neg, \wedge\}$-Fragment der intuitionistischen Sprache $\mathcal{L}_i$. Nach dem vorigen Korollar ist eine Formel in diesem Fragment genau dann ein $\mathbf{IL}$-Theorem, wenn sie ein $\mathbf{KL}$-Theorem ist. ∎

Einerseits entsteht die klassische Logik aus der intuitionistischen Logik indem wir zu den Heyting-Axiomen beispielsweise das Schema DNB oder das Peirce-Schema hinzufügen. In diesem Sinne ist intuitionistische Logik ein echter Teil der klassischen Logik. Andererseits ermöglicht das Korollar aber auch, **KL** als Teil von **IL** zu betrachten, nämlich als das Fragment von **IL** in einer Sprache, die sich auf die Junktoren $\neg$ und $\wedge$ beschränkt. Die Klassische Logik wäre demnach nichts anderes als die intuitionistische Logik einer kleineren Sprache; vgl. dazu unsere Anmerkung auf p. 177. Richtig verstanden, d.h. übersetzt, meint der klassische Logiker mit $\neg\neg A \to A$ nichts anderes als $\neg(\neg\neg A \wedge \neg A)$. Das ist eine Instanz des Prinzips vom ausgeschlossenen Widerspruch – nichts also, was der Intuitionist zurückweisen möchte. Der vordergründige Disput über das richtige Verständnis der Junktoren wird so sytematisch ausgeräumt. Das Korollar präsentiert die intuitionistische Position als Resultat eines Sprachwechsels – genauer, einer Spracherweiterung. Ganz im Sinne von Quine [236]: “Change of logic, change of subject.”

15. Funktionale und Umgebungsrahmen

In diesem abschließenden Abschnitt betrachten wir sehr kurz einige mehr oder weniger naheliegende Variationen und Alternativen zur Modellierung modaler Operatoren in Kripke-Rahmen.

Funktionale Kripke-Rahmen. Was geschieht, wenn die Relation R in einem Rahmen (W, R) *funktional* ist, also die Bedingung

$$\text{(Fun)} \qquad Rab \;\&\; Rac \;\Rightarrow\; b = c$$

für alle Punkte erfüllt ist? Wir fragen hier nach *funktionalen Kripke-Rahmen*, d.h. wenn wir Rab ersetzen durch die Notation $f(a) = b$, dann betrachten wir nun Rahmen des Typs

$$(W, f) \text{ mit } f : W \longrightarrow W$$

und der Modellbedingung

$$(\Box f) \qquad a \models \Box A \text{ gdw } f(a) \in W \;\Rightarrow\; f(a) \models A.$$

Zunächst fällt auf, daß $\Box$ nicht mehr als Allquantor interpretiert wird, sondern als eine Indexverlagerung: $\Box A$ am Index a fordert dazu auf, A am Index $f(a)$ zu betrachten, falls dieser in W existiert, d.h. falls f für a definiert ist. Da alle Indizes negationsvollständig sind, so ist nun $\Box A \vee \Box\neg A$ wahr an allen Punkten, für die f definiert ist; an allen anderen Punkten gilt das Schema auf triviale Weise. Äquivalent ausgedrückt, impliziert in diesen Rahmen Möglichkeit Notwendigkeit:

D^c. $$\Diamond A \to \Box A.$$

Der vollständige Kollaps der Modalitäten wird erreicht, wenn (Fun) mit der Bedingung der Erweiterbarkeit (Erw) kombiniert wird: $\forall a \exists b Rab$. Das ist die Annahme, daß für jeden Punkt a, $f(a)$ in W ist, weshalb wir zur einfacheren Modellbedingung

($\Box f$) $$a \models \Box A \text{ gdw } f(a) \models A$$

übergehen können. Nun gilt auch die Umkehrung von D^c, d.h. wir haben

D! $$\Diamond A \leftrightarrow \Box A.$$

Wir haben oben schon bewiesen, daß die Klasse der Kripke-Rahmen, in denen R erweiterbar – bzw. f total – ist, die Logik **KD** festlegt. Da die Funktionalität (FUN) von R und das Schema D^c ein korrespondierendes Paar bilden, so ergibt sich **KD**! als die minimale Logik funktionaler Kripke-Rahmen (W, f), in denen f überall definiert ist.

Transformationsrahmen. Die Notwendigkeit einer Aussage könnten wir auch so erklären: Gleichgültig, was an unserer Welt verändert werden mag, die Aussage bleibt wahr. Die Veränderung einer Welt stellen wir durch eine Funktion $f : W \longrightarrow W$ dar. Solche Funktionen sammeln wir in einer Menge F. Die gerade gegebene Erklärung quantifiziert nicht unmittelbar über eine Menge W von Welten, sondern über eine (nichtleere) Menge F von Veränderungsfunktionen:

($\Box F$) $$a \models \Box A \text{ gdw } \forall f \in F : \ f(a) \models A.$$

Unter der gewöhnlichen Definition $\Diamond = \neg\Box\neg$ erhalten wir ferner

($\Diamond F$) $$a \models \Diamond A \text{ gdw } \exists f \in F : \ f(a) \models A.$$

Wir nehmen hier an, daß alle Veränderungsfunktionen für alle Punkte in W definiert sind. Das entspricht der Bedingung der Erweiterbarkeit der Relation R. So ist einerseits die Gültigkeit von D, $\Box A \to \Diamond A$, wie im Falle der funktionalen Kripke-Rahmen garantiert. Andererseits zeigen die Klauseln für $\Box$ und $\Diamond$ diese Operatoren wieder als distinkte Quantoren, was den trivialisierenden Kollaps D^c verhindert. Die Klasse aller Transformationsrahmen (W, F) stellt sich tatsächlich als eine alternative Modellierung des Basissystems deontischer Logik, **KD**, dar. Transformationsrahmen wurden erstmals von van Fraassen [78] und Garson [104] vorgestellt; eine ausführliche Diskussion findet sich in [28] und [144, §2.10] .

Umgebungsrahmen. In funktionalen Kripke-Rahmen betrachten wir Funktionen

$$W \longrightarrow W,$$

in allgemeinen Kripke-Rahmen solche des Typs

$$W \longrightarrow \wp(W).$$

Die nächste Sprosse auf dieser Leiter sind Rahmen mit Funktionen

$$W \longrightarrow \wp(\wp(W)).$$

Eine solche Progression von Funktionstypen ist natürlich völlig nichtssagend, ohne eine Erklärung, wie auch dieser letzte Schritt für die Interpretation modaler Formeln genutzt werden kann. Dazu erinnern wir an eine nützliche Terminologie: Eine *Proposition* ist eine Menge von Punkten. Jeder Formel A entspricht in einem gegeben Modell eine Proposition $[\![A]\!] = \{a \in W : a \models A\}$. Wann sollten wir sagen, daß ein Punkt a zur Proposition $[\![\Box A]\!]$ gehört? Eine sicher richtige Antwort ist diese:

> $a \in [\![\Box A]\!]$ gdw $[\![A]\!]$ zu den Propositionen gehört, die am Punkt a notwendig sind.

Diese Antwort können wir modellieren, indem wir uns jeden Punkt $a \in W$ umgeben denken von Propositionen, die am Punkt a notwendig sind. Eine solche Umgebung $N(a)$ notwendiger Propositionen ist eine Menge in $\wp(\wp(W))$. So wird aus der sicher richtigen Antwort diese Modellbedingung

$$(\Box N) \qquad a \models \Box A \text{ gdw } [\![A]\!] \in N(a)$$

für sogenannte *Umgebungsrahmen* – auch Nachbarschaftsrahmen genannt – (W, N) mit $N : W \longrightarrow \wp(\wp(W))$. Aus der Definition $\Diamond = \neg\Box\neg$ ergibt sich

$$(\Diamond N) \qquad a \models \Diamond A \text{ gdw } W \setminus [\![A]\!] \notin N(a).$$

Umgebungsrahmen gehen auf Dana Scott und Richard Montague zurück. Eine erste systematische Untersuchung solcher Rahmen ist Segerbergs Monographie [260]; eine neuere Behandlung ist [218]. Mit Umgebungsrahmen wird die Modellierungsschwelle von Kripke-Rahmen deutlich unterschritten. Damit ist gemeint, daß es Klassen von Umgebungsrahmen gibt, deren Theorien jeweils Logiken sind, die schwächer sind als die kleinste normale Modallogik **K**.

Betrachten wir die für **K** charakteristische Regel

$$\text{RK.} \qquad \frac{A_1 \wedge \cdots \wedge A_n \to B}{\Box A_1 \wedge \cdots \wedge \Box A_n \to \Box B}\ (n \geq 0)$$

und ihre Fälle RN ($n = 0$), RM ($n = 1$) und RC ($n = 2$).

Beobachtung 90. *Keine der Regeln RN, RM und RC ist gültig in der Klasse $\mathbb{U}$ aller Umgebungsrahmen.*

Beweis. *Ad* RC: In einem Gegenmodell gilt (nach atomaren Einsetzungen für A_1, A_2 und B) $[\![A_1]\!] \cap [\![A_2]\!] \subseteq [\![B]\!]$ und es gibt einen Punkt a mit (a) $[\![A_1]\!]$ und $[\![A_2]\!]$ in $N(a)$, während $[\![B]\!] \notin N(a)$.

Beispiel: $W{=}\{a_1, a_2, b\}$, $N(a_1){=}\{\{a_1, b\}, \{a_2, b\}\}$; $[\![A_1]\!]{=}\{a_1, b\}$, $[\![A_2]\!]{=}\{a_2, b\}$, $[\![B]\!]{=}\{b\}$.

Rahmen, die unter (endlichen) Schnitten und unter Obermengen abgeschlossen sind, lassen ein solches Gegenmodell nicht zu.

Ad RM: In einem Gegenmodell gilt $[\![A]\!] \subseteq [\![B]\!]$ und es gibt einen Punkt a mit $[\![A]\!] \in N(a)$, während $[\![B]\!] \notin N(a)$.

Beispiel: $W{=}\{a, b\}$, $N(a){=}\{\{a\}\}$; $[\![A]\!]{=}\{a\}, [\![B]\!]{=}\{a, b\}$.

In Rahmen, die unter Obermengen abgeschlossen sind, ist ein solches Gegenmodell nicht möglich.

Ad RN: Jeder Rahmen mit einem Punkt a so, daß $W \notin N(a)$ erzeugt Gegenmodelle. ∎

Wo liegt nun die unterste Schwelle der Modellierung, d.h. welche Logik ist die Theorie der Klasse $\mathbb{U}$ *aller* Umgebungsrahmen? Das ist die Logik **E**, bestehend aus allen Tautologien, sowie den Regeln MP und

RE. $$\frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B}\,.$$

Einen Beweis, daß $\mathbf{E} = Th(\mathbb{U})$, enthält [260] und das zugänglichere Buch [53]. Modallogiken, die unter RE abgeschlossen sind, heißen *kongruent* oder *klassisch* (wie in Chellas [53]).

Ersichtlich ist die Wahrheitsbedingung für $\Box A$ in Umgebungsmodellen nun sehr nahe an der Formel selbst: $a \models \Box A$ wird "übersetzt" zu $[\![A]\!] \in N(a)$ und umgekehrt. Das macht auch die korrespondierende Paarung von Schemata und Rahmenbedingungen geradezu trivial. So korrespondiert zum Beispiel das Schema

R. $$\Box A \wedge \Box B \leftrightarrow \Box(A \wedge B)$$

zur Bedingung

(*reg*) $X \in N(a)$ und $Y \in N(a)$ gdw $X \cap Y \in N(a)$,

und das zur RN-Regel äquivalente Schema

N. $$\Box\top$$

entspricht – wie im Beweis der letzten Beobachtung zu sehen war – der Bedingung

$$(nec) \qquad W \in N(a).$$

Wir wissen, daß **E** zusammen mit R und N eine Axiomatisierung von **K** darstellt. Also ...

> ... ist **K** die Theorie der Umgebungsrahmen, die unter den Bedingungen (*reg*) und (*nec*) stehen.

Nun haben wir zwei Modellierungen von **K**: die eine mit Kripke-Rahmen, die andere mit Umgebungsrahmen. In welcher Beziehung stehen diese beiden Arten von Rahmen zueinander? Wir werden zeigen, daß man Kripke-Rahmen als eine bestimmte, "normale" Art von Umgebungsrahmen auffassen kann, und umgekehrt.

Eine Umgebung $N(a)$ $(a \in W)$ wollen wir *normal* nennen,[38] wenn

$$(norm) \qquad X \in N(a) \text{ gdw } \bigcap N(a) \subseteq X.$$

Normale Umgebungsrahmen sind solche, in denen alle Punktumgebungen normal sind.

SATZ 91. *Zu jedem Kripke-Rahmen gibt es einen punktweise äquivalenten normalen Umgebungsrahmen; zu jedem normalen Umgebungsrahmen gibt es einen punktweise äquivalenten Kripke-Rahmen.*

BEWEIS. Sei (W, R) ein Kripke-Rahmen. Wir definieren (für alle $a \in W, X \subseteq W$):

$$(*) \qquad X \in N^R(a) \text{ gdw } R(a) \subseteq X.$$

Dann ist $\bigcap N^R(a) = R(a)$ und also ist $N^R(a)$ normal, für jeden Punkt a in W, d.h., (W, N^R) ist ein normaler Umgebungsrahmen. Wir zeigen nun, daß für jede Interpretation I, (W, R, I) und (W, N^R, I) punktweise äquivalent sind, d.h. für alle Formeln A,

$$a \models A \text{ in } (W, R, I) \text{ gdw } a \models A \text{ in } (W, N^R, I).$$

Der einzige interessante Fall in der Induktion ist $A = \Box B$:

$$\begin{aligned} a \models \Box B \text{ in } (W, R, I) &\text{ gdw } \forall b : Rab \Rightarrow b \models B \\ &\text{ gdw } R(a) \subseteq [\![B]\!] \\ &\text{ gdw } [\![B]\!] \in N^R(a) && \text{Def. } (*) \\ &\text{ gdw } a \models \Box B \text{ in } (W, N^R, I) \end{aligned}$$

[38] Engl. *augmented*.

Für die Umkehrung sei (W, N) ein normaler Umgebungsrahmen. Wir definieren

$$(\dagger) \qquad R^N ab \text{ gdw } b \in \bigcap N(a),$$

d.h. $R^N(a) = \bigcap N(a)$. Dann gilt für alle Formeln A unter einer beliebigen Interpretation I,

$$a \models A \text{ in } (W, N, I) \text{ gdw } a \models A \text{ in } (W, R^N, I).$$

Hier wieder der Fall $A = \Box B$:

$$\begin{aligned} a \models \Box B \text{ in } (W, N, I) &\text{ gdw } [\![B]\!] \in N(a) \\ &\text{ gdw } \textstyle\bigcap N(a) \subseteq [\![B]\!] && (norm) \\ &\text{ gdw } R^N(a) \subseteq [\![B]\!] && \text{Def. } (\dagger) \\ &\text{ gdw } \forall b : R^N ab \Rightarrow b \models B \text{ in } (W, R, I) \end{aligned}$$

■

KOROLLAR 92. $\mathbf{K} = Th(Norm(\mathbb{U}))$.

BEWEIS. Für die $\subseteq$-Richtung genügt es, sich von der Gültigkeit aller Axiome und Regeln von **K** in normalen Umgebungsmodellen zu überzeugen. Dazu gehen wir hier von folgender Axiomatisierung aus: Die Menge der Tautologien in der modalen Sprache, vermehrt um die Schemata R und N, und abgeschlossen unter den Regeln RE und MP. Die Regeln sind schnell verifiziert. Die Schemata R und N korrespondieren jeweils zu (*reg*) und (*nec*). Es genügt also zu zeigen, daß die für normale Umgebungen charakteristische Bedingung (*norm*) diese beiden Bedingungen impliziert, was trivial ist.

Für die umgekehrte Richtung nehmen wir an, daß $A \notin \mathbf{K}$. Dann gibt es ein A-falsifizierendes Kripke Modell $\mathcal{M}$. Nach dem gerade bewiesenen Satz gibt es ein zu $\mathcal{M}$ punktweise äquivalentes (und normales) Umgebungsmodell; also $A \notin Th(Norm(\mathbb{U}))$. ■

Die Bedingungen (*nec*), (*reg*) und (*norm*) für Umgebungsrahmen lassen sich in Beziehung setzen zum Begriff eines Mengenfilters. Ein *Filter* über einer Menge W ist eine nichtleere Menge $F \subseteq \wp(W)$ so, daß

$$(\text{F1}) \qquad X, Y \in F \Rightarrow X \cap Y \in F,$$

$$(\text{F2}) \qquad X \in F \text{ und } X \subseteq Y \Rightarrow Y \in F.$$

Die Bedingung (F1) ist nichts anderes als die $\Rightarrow$-Richtung von (*reg*). Die andere Richtung ist äquivalent zu (F2). (Denn, angenommen $X \cap Y \in F$.

Dann $X \in F$ nach (F2), da $X \cap Y \subseteq X$. Umgekehrt nehmen wir an, $X \in F$ und $X \subseteq Y$. Dann $X = X \cap Y$ und also $Y \in F$ nach (*reg*$\Leftarrow$).) Schließlich folgt (*nec*) unmittelbar aus (F2). Mit anderen Worten, eine Umgebung $N(a)$ steht genau dann unter den Bedingungen (*nec*) und (*reg*), wenn $N(a)$ ein Filter ist. Da **K** = **E**+N+R und (N,*nec*) sowie (R,*reg*) jeweils korrespondierende Paare sind, so haben wir – wie auf p. 195 schon in anderen Worten erwähnt – neben Korollar 92 ein weiteres Richtigkeits- und Vollständigkeitsresultat für **K**:

> **K** ist die Theorie der Klasse der Umgebungsrahmen, in denen alle Umgebungen Filter sind.

Man beachte, daß diese Klasse nicht mit $Norm(\mathbb{U})$ koinzidiert. Von normalen Umgebungen haben wir im Beweis des letzten Korollars gesehen, daß sie die Bedingungen (*nec*) und (*reg*) erfüllen; also sind normale Umgebungen Filter. Aber nicht jeder Filter $N(a)$ ist eine normale Umgebung. Letztere sind dadurch ausgezeichnet, daß sie ihren Kern $\bigcap N(a)$ enthalten – oder, äquivalent (gegeben F2): sie sind unter beliebigen, statt nur endlichen Schnitten abgeschlossen.

Wir haben oben beobachtet, daß für Umgebungsrahmen die korrespondierende Paarung von Formelschemata und Rahmenbedingungen besonders einfach ist. Das legt den Gedanken nahe, daß auf der Basis der Richtigkeit und Vollständigkeit des Systems **E** entsprechende Resultate für jede schematische Erweiterung von **E** auf nahezu triviale Weise garantiert sind. Das ist jedoch keineswegs so. Die Logiken von Thomason [280] und Fine [75], die bezüglich Kripke-Rahmen unvollständig sind, sind es ebenfalls in Bezug auf Umgebungssrahmen (Gerson [107]). Shehtman [264] gibt ein weiteres Beispiel einer klassischen Modallogik, die nicht die Theorie einer Klasse von Umgebungsrahmen ist; siehe auch [265].

IV.

KONDITIONALE

1. Verschiedene Arten von "Wenn ..., dann ..."

Konditionalsätze, kurz *Konditionale*, sind Sätze der Form

(1) Wenn A, dann B.

Den Wenn-Teil des Satzes nennen wir das *Antezedens*, den Dann-Teil das *Konsequens* des Konditionals. Konditionale können grammatisch in einem indikativischen ("realis") oder einem konjunktivischen ("irrealis", engl. *subjunctive*) Modus stehen:

(Ind) *Realis*: Wenn A der Fall ist, dann ist B der Fall.
(Konj) *Irrealis*: Wenn A der Fall wäre, dann wäre B der Fall.

Daß der grammatische Unterschied zugleich einen semantischen markiert, sieht man sehr schnell an Paaren wie den folgenden:

(2) Wenn die Suppe versalzen ist, dann habe ich es nicht bemerkt.
(3) Wenn die Suppe versalzen wäre, dann würde ich es nicht bemerken.

(4) Wenn er ihr bescheid gesagt hat, dann hat sie es nicht gehört.
(5) Wenn er ihr bescheid gesagt hätte, dann würde sie es nicht gehört haben.

(6) Wenn Aigner nicht das Tor geschossen hat, dann war es Meier.
(7) Wenn Aigner nicht das Tor geschossen hätte, dann wäre es Meier gewesen.

Die beiden Sätze eines Paares unterscheiden sich nur im Modus. Zu jedem Paar läßt sich mühelos eine passende Geschichte – ein Kontext – denken, so daß der erste Satz wahr und der zweite Satz falsch ist. (Man stelle sich vor, die Suppe sei nicht versalzen, oder ich habe nichts bemerkt; oder er habe nicht Bescheid gegeben, oder sie habe nichts gehört; oder einer von beiden habe ein Tor geschossen.) Also haben Konditionale im Indikativ im allgemeinen andere Wahrheitsbedingungen als Konditionale im Konjunktiv.

Für die Wahrheitsbedingung eines indikativischen Konditionals (Ind) haben wir eine einfache Kandidatin: "Wenn A, dann B" ist genau dann wahr, wenn die materiale Implikation, $A \to B$ wahr ist, d.h. wenn A falsch oder B wahr ist. Diese Position ist nicht unmittelbar überzeugend. Ist zum Beispiel der Satz (6) schon allein deshalb wahr, weil Aigner ein Tor geschossen hat? Offensichtlich müssen wir mehr sagen, um die Identifizierung des materialen mit dem indikativischen Konditional auch nur annähernd plausibel zu machen. So müssen wir insbesondere erklären, warum die sogenannten Paradoxien der materialen Implikation der Identifizierung nicht im Wege stehen – d.h. eine Theorie anbieten, die erklärt warum wir uns irren, wenn wir dies glauben, und möglichst auch, warum dieser Irrtum so naheliegend ist. Auf eine solche Irrtumstheorie werden wir zurückkommen.

Der Vorschlag, das konjunktivische mit dem materialen Konditional zu identifizieren, ist sicher aussichtslos. Das materiale Konditional wird nur unter einer Bedingung falsch: Wenn das Antezedens wahr und das Konsequens falsch ist. Betrachten wir jedoch die Sätze (3), (5) und (7) im Kontext der Geschichten, die wir uns soeben dazu gedacht haben, dann haben wir diese aus einem anderen Grunde für falsch befunden, als daß es sich tatsächlich so verhielte, daß das Antezedens wahr und das Konsequens falsch wäre. Ob (3) falsch ist, hängt nicht davon ab, ob die Suppe versalzen *ist*.

Es gibt einen weiteren Grund, warum Konditionale der Form (Konj) im allgemeinen keine materialen Implikationen sein können. Letzere sind schon dann wahr, wenn das Antezedens falsch ist. Konditionale im Irrealis werden aber typischerweise gerade dann gebraucht, wenn der Sprecher davon ausgeht, daß das Antezedens tatsächlich falsch ist. Wären diese Konditionale im materialen Sinne zu verstehen, dann wären sie unterschiedslos wahr. D.h. konjunktivische Konditionale könnten im typischen Fall keine Informationen vermitteln. Wenn die Suppe nicht versalzen ist, dann wäre unter der Annahme, daß konjunktivische Konditionale materiale Implikationen wiedergeben,

(8) Wenn die Suppe versalzen wäre, dann würde ich es nicht bemerken

genauso wahr wie

(9) Wenn die Suppe versalzen wäre, dann würde ich es bemerken.

Aber im gedachten Kontext sehen wir die Sache sicher anders: (8) ist falsch und (9) ist wahr.

Jacksons Tatächlichkeitsargument. Wenn konjunktivische Konditionale keine materialen sind, was für Konditionale sind es dann? Betrachten wir das folgende Paar:

(10) Wenn Peter zwei Meter groß wäre, dann wäre er größer als er es tatsächlich ist;

(11) ?Wenn Peter zwei Meter groß ist, dann ist er größer als er es tatsächlich ist.

Im Gegensatz zum ersten, konjunktivischen Konditional, ist das zweite, im Indikativ irgendwie verunglückt. (Wir deuten das hier und im folgenden durch ein hochgestelltes Fragezeichen an.) Das Unglück ist kein grammatisches, denn (11) ist syntaktisch richtig gebildet. Der Satz (11) ist in einer Weise fehlerhaft, die jedem Sprecher gleich auffällt – auch wenn dieser nicht sogleich sagen kann, welche Regel hier verletzt ist.

Ganz allgemein ist es so, daß bei allen Paaren der Form

(12) Wenn ... der Fall wäre, dann wäre einiges anders als es (tatsächlich) ist.

(13) ?Wenn ... der Fall ist, dann ist einiges anders als es (tatsächlich) ist.

das erste, konjunktivische Konditional sinnvoll gebraucht werden kann, während man sich für das zweite, indikativische Konditional keinen sinnvollen Gebrauch vorstellen kann. Woran liegt das? Nehmen wir einmal an, Peter behauptet:

(14) Ich reiche nicht an die Lampe heran.

Wenn Peter jetzt weiter behauptet:

(15) Wenn ich größer als 2 Meter wäre, dann würde ich an die Lampe heranreichen,

dann behauptet er mit (15) auch

(16) Wenn ich größer als 2 Meter wäre, dann wäre einiges anders als es (tatsächlich) ist.

Wenn Peter jedoch nach (14) behauptet:

(17) Wenn ich größer als 2 Meter bin, dann reiche ich an die Lampe heran,

dann will er damit sicher nicht auch sagen:

(18) ?Wenn ich größer als 2 Meter bin, dann ist einiges anders als es (tatsächlich) ist.

Der letzte Satz ist allenfalls eine merkwürdige Weise, die Falschheit des Antezedens auszudrücken. Das tut (im typischen Fall) auch (16) – aber eben auf eine Weise, die überhaupt nicht merkwürdig ist. Das Antezedens im Konjunktiv lädt zur Betrachtung möglicher Welten ein, in denen die Dinge anders sein können als sie es tatsächlich, d.h. in der aktualen Welt, sind. In einer solchen kontrafaktischen Welt ist es wahr, daß in ihr einiges anders ist, als in der Welt, auf die das “tatsächlich” verweist. Würde das Antezedens im Indikativ ebenfalls zur Betrachtung möglicher Welten einladen, dann müßte es im Prinzip ebenfalls möglich sein, mit dem Konsequens etwas Wahres zu sagen, nämlich daß einiges anders ist, als es ist. Aber das Konsequens von (18) kann nicht wahr sein: Nichts ist anders, als es ist. Also lädt das Antezedens eines indikativischen Konditionals nicht zur Betrachtung bloß möglicher Welten ein.

In Sätzen nach dem Muster (12) verweist das “tatsächlich” im Konsequens zurück auf die aktuale Welt, nachdem durch das Antezedens mögliche Welten in den Focus geraten sind. Humberstone [143, pp. 930f.] weist darauf hin, daß das “tatsächlich” diese Rolle auch im Antezedens spielen kann:

(19) Wenn Peter größer wäre, als er es (tatsächlich) ist, dann würde er an die Lampe heranreichen;

(20) $^{?}$Wenn Peter größer ist, als er es (tatsächlich) ist, dann reicht er an die Lampe heran.

Wieder hat das Antezedens des indikativischen Gegenstücks (20) zu (19) keine Chance wahr zu sein. Die naheliegende Erklärung ist: Dem konjunktivischen Konditional (19) gelingt es im Antezedens alternative Möglichkeiten ins Auge zu fassen, welche dann im Hinblick auf das Konsequens untersucht werden sollen. Dem Konditional (20) im Indikativ gelingt das nicht; hier wird die aktuale Welt gar nicht verlassen. Die indikativische Konstruktion ist nicht dazu geeignet, alternative Möglichkeiten ins Spiel zu bringen.

Jackson [147, p. 129] faßt das so zusammen: Konjunktivische Konditionale sind mögliche Welten-Konditionale, indikativische sind es nicht. Der erste Teil dieser These findet bei Philosophen weitgehend Zustimmung. Der zweite Teil ist umstritten. Wir werden später (im Abschnitt über indikativische Konditionale) auf Jacksons Tatsächlichkeitsargument zurückkommen.

Konjunktiv und irrealis. Die Unterscheidung zwischen Konditionalen im Indikativ und solchen im Konjunktiv ist eine grammatische Unterscheidung, welche – so haben wir gesehen – einen wichtigen semantischen Unterschied markiert. Viele Autoren in der Tradition der Philosophischen

Logik ziehen es vor, die indikativischen Konditionalsätze den *kontrafaktischen* gegenüberzustellen. Grob gesagt, sind die konjunktivischen Konditionalsätze diejenigen, in denen das Antezedens typischerweise kontrafaktisch ("irreal") zu verstehen ist. Auf den ersten Blick ist das eine merkwürdige Kontrastierung, denn "indikativisch" bezeichnet eine syntaktische Eigenschaft, während "kontrafaktisch" eine semantische Eigenschaft bezeichnet. Es ist nun einerseits tatsächlich so, wie wir gerade gesehen haben, daß ein Sprecher einem Hörer ein kontrafaktisches Antezedens nur mit einem konjunktivischen, nicht mit einem indikativischen Konditional vor Augen stellen kann. Andererseits gilt aber nicht umgekehrt, daß jeder gute Gebrauch eines konjunktivischen Konditionals die Falschheit des Antezedenz unterstellt. Man betrachte:

(21) Wenn Peter die Pizza holen würde, dann hätte Ulla Zeit zum Baden;
(22) Wenn Peter sich entschuldigen würde, dann würde Ulla ihm sicher verzeihen.

Kein Sprecher muß diese Sätze zurückziehen, wenn Peter tatsächlich eine Pizza holt bzw. sich entschuldigt. Das Tempus scheint hier eine Rolle zu spielen, denn Äußerungen der in die Vergangenheit transponierten Sätze

(23) Wenn Peter die Pizza geholt hätte, dann hätte Ulla Zeit zum Baden gehabt;
(24) Wenn Peter sich entschuldigt hätte, dann würde Ulla ihm sicher verziehen haben,

geben wohl typischerweise zu verstehen, daß Peter keine Pizza geholt bzw. sich nicht entschuldigt hat.

Die syntaktische Markierung semantischer Unterscheidungen ist offenbar auch im bereits eingeschränkten Bereich der konjunktivischen Konditionalsätze eine recht komplexe Angelegenheit. Wir werden hier keine befriedigende Klärung dieser Frage herbeiführen können. Wir werden einfach diejenigen konjunktivischen Konditionale, deren Antezedens in typischen Kontexten zur Betrachtung kontrafaktischer Möglichkeiten auffordert, "kontrafaktische Konditionale" nennen. Für diese Konditionale werden wir eine semantische Theorie angeben. Damit mag es zwar für manche Sätze offenbleiben, ob sie in den Anwendungsbereich der Theorie fallen. Aber welche konjunktivischen Formen in typischen Kontexten paradigmatische Fälle kontrafaktischer Konditionale sind, dürfte hinreichend deutlich sein.[1]

[1] Das Reich der Konditionalsätze ist von schwer zu ordnender Vielfalt. Hier sind einige weitere Konditionale:

> Selbst wenn er es getan hätte, so könnte er nicht belangt werden.

2. Strikte Konditionale

Ein Satz wie

(25) Wenn Peter eingeladen wäre, würde er jetzt hier sein

ist von der Form

Wenn A der Fall wäre, dann würde B der Fall sein.

Solche Sätze, d.h. kontrafaktische Konditionale, wollen wir kurz so notieren: [2]

$$A \sqsupset B.$$

Wir wissen, daß es für die Beurteilung von $A \sqsupset B$ nicht ausreicht, die Wahrheitswerte von A und B in der aktualen Welt zu betrachten. Andere mögliche Welten müssen ebenfalls in den Blick genommen werden. Angenommen (25) ist wahr. Dann ist es sicher nicht logisch zwingend, daß Peter in jeder Welt, in der er eingeladen ist, der Einladung auch folgt. Das Konditional (25) fordert nicht, daß es keine logische Möglichkeit gibt, in der A wahr und B falsch ist. Aber, so könnte man vorschlagen, das Konditional behauptet, daß es in einem eingeschränkten Sinn von Möglichkeit, keine Möglichkeit gibt, daß A wahr und B falsch ist: In allen als naheliegend zu betrachtenden ("plausiblen") Möglichkeiten ist $A \wedge \neg B$ nicht der Fall. Wenn wir eine Relation R zwischen jeweils zwei Welten so interpretieren:

Rab gdw b, von a aus betrachtet, eine naheliegende Möglichkeit ist,

dann können wir die soeben vorgeschlagene schematische Wahrheitsbedingung für $A \sqsupset B$ so festhalten:

(26) $a \models A \sqsupset B$ gdw $\forall b$: wenn Rab und $b \models A$, dann $b \models B$.

(Hier deutet $\models$ die Wahrmacherrelation in Modellen an: Für $a \models A$ lies "A ist in der Welt a wahr.").

Nun ist die rechte Seite von (26) nichts anderes als die Wahrheitsbedingung für die materiale Implikation im Skopus eines Notwendigkeitsoperators, welchen wir hier versuchsweise als Plausibilitätsoperator interpretieren wollen:

Selbst wenn er es getan hat, so kann er nicht belangt werden.
Kekse sind auf dem Tisch, wenn Du welche möchtest.
Wenn Peter Klassensprecher wird, dann wird Claudia Papst.

Wir werden uns hier mit solchen Konditionalen nicht beschäftigen und verweisen auf die einschlägige Literatur, z.B. [25] und [143].

[2] Diese Notation geht auf Segerberg zurück. Andere Autoren benutzen das von Lewis [182] eingeführte Symbol $\Box\!\!\rightarrow$.

$$a \models \Box(A \to B) \text{ gdw } \forall b\text{: wenn } Rab \text{ dann } b \models A \to B.$$

Die materiale Implikation im Skopus eines solchen Operators nennt man auch *strikte Implikation* (nach C.I. Lewis). Diese können wir durch eine neue, definierte zweistellige Verknüpfung wiedergeben:

$$A \twoheadrightarrow B := \Box(A \to B).$$

Der hier betrachtete Vorschlag lautet also: Das kontrafaktische Konditional $\sqsupset$ hat die Wahrheitsbedingung der strikten Implikation $\twoheadrightarrow$, wobei wir die Zugangsrelation in der oben angedeuteten Weise interpretieren wollen. In der Frühzeit der Behandlung kontrafaktischer Konditionalsätze wurde dieser Vorschlag gelegentlich auf die eine oder andere Weise gemacht. Es läßt sich recht schnell zeigen, daß dies ein falscher Weg ist. Wir nutzen die Gelegenheit, um ein wenig die logische Struktur strikter Implikationen zu erkunden. Schließlich handelt es sich bei strikten Implikationen um eine wichtige Klasse von indikativischen Konditionalen, die wir bisher noch gar nicht erwähnt haben.

Die Semantik strikter Konditionale. Wir betrachten eine Sprache mit den üblichen wahrheitsfunktionalen Junktoren, welche um den Junktor $\twoheadrightarrow$ (strikte Implikation) erweitert ist. Eine solche Sprache interpretieren wir in Modellen

$$(W, R, I)$$

auf Kripke-Rahmen (W, R).

Erinnerung:[3] Ein *Kripke-Rahmen* besteht aus einer nichtleeren Menge W (Punkte, Indizes, "mögliche Welten") sowie einer Relation $R \subseteq W \times W$ (die Zugangsrelation zwischen Punkten). Auf einem Rahmen (W, R) definieren wir ein *Modell* (W, R, I), indem wir eine *Interpretation* I angeben, welche jedes Atom der Sprache an einem Punkt auf genau einen der Werte 0 oder 1 abbildet. Eine *Wahrmacherrelation* $\models$ baut auf der Interpretation I der Atome auf und erlaubt rekursiv den Wahrheitswert einer jeden Formel an jedem Punkt zu bestimmen. Die Rekursion für $\models$ beginnt mit $a \models P$ gdw $I(P, a) = 1$ (für Atome P) und setzt sich dann für wahrheitsfunktionale Zusammensetzungen in üblicher Weise fort. Für strikte Implikationen haben wir die oben bereits eingeführte Bedingung $(\twoheadrightarrow)$. — Nun sagen wir, daß eine Formel B aus einer (möglicherweise

[3] Vgl. das Kapitel III.5 über Modallogik.

leeren) Menge X von Annahmen logisch genau dann *folgt* ($X \models B$), wenn an jedem Punkte a eines (beliebigen) Modells gilt: Wenn $a \models A$ (für jede Formel $A \in X$), dann $a \models B$.[4]

(*Ende der Erinnerung.*)

Die Wahrheitsbedingung ($\twoheadrightarrow$) gibt die allgemeinste semantische Definition der strikten Implikation ab. Wir können, darauf aufbauend, den Bereich strikter Implikationen differenzieren, indem wir die Relation R unter Bedingungen stellen. So schränken wir die Menge der zu betrachtenden Rahmen und also Modelle ein, was dazu führt, daß mehr Paare (X, A) als logische Folgerungen ausgezeichnet werden. Was unsere Hypothese, kontrafaktische seien strikte Konditionale, zu Fall bringt, ist jedoch nicht der Umstand, daß sie *zu wenige* Folgerungen generiert, sondern – umgekehrt – daß sie ohne weiteres Folgerungen als logische auszeichnet, die es unter der beabsichtigten Interpretation sicher nicht sind. Die Relation R unter weitere Bedingungen zu stellen, kann die Sache daher nicht besser machen. Hier sind einige Folgerungen, die für jede Art strikter Implikation, jedoch nicht für kontrafaktische Konditionale gelten:

BEOBACHTUNG 1. *In beliebigen Klassen von Kripke-Rahmen gilt:*
1. $A \twoheadrightarrow B, B \twoheadrightarrow C \models A \twoheadrightarrow C$ *(Trans.itivität);*
2. $A \twoheadrightarrow B \models A \wedge C \twoheadrightarrow B$ *(Verst.ärkung des Antezedens);*
3. $A \twoheadrightarrow B \models \neg B \twoheadrightarrow \neg A$ *(Kontrap.osition).*

BEWEIS. *Ad* 1: Wir nehmen an, daß

$$(1)\ a \models A \twoheadrightarrow B \quad \text{und} \quad (2)\ a \models B \twoheadrightarrow C.$$

Zz. $a \models A \twoheadrightarrow C$. Also nehmen wir ferner an, daß

$$(3)\ Rab \text{ und } b \models A.$$

(Zz. $b \models C$.) Aus (1) und (2) folgt jeweils:

$$(1') \qquad \forall x : Rax \;\&\; x \models A \;\Rightarrow\; x \models B,$$

$$(2') \qquad \forall y : Ray \;\&\; y \models B \;\Rightarrow\; y \models C.$$

Aus (1′) und (3) folgt (4) $b \models B$, woraus mit (2′) und (3) die gewünschte Konklusion $b \models C$ folgt.

Ad 2 und 3: Ähnlich. ∎

[4] Die Relation der Folgerung ist natürlich relativ zu einer Klasse von Rahmen, auf denen die betrachteten Modelle aufsitzen. Der Einfachheit halber möge der Leser hier an die Klasse *aller* Rahmen denken.

Nun gehören aber (Trans), (Verst) und (Kontrap) zu den auffälligsten Folgerungsverhältnissen, die wir von kontrafaktischen Konditionalen *nicht* erwarten. Die folgenden, intuitiv ungültigen Folgerungen sind Instanzen dieser Prinzipien:

(27) Wenn der Kamin brennen würde, dann wäre es gemütlich.
Wenn das Haus brennen würde, dann würde der Kamin brennen.
$^{?}$Also: Wenn das Haus brennen würde, dann wäre es gemütlich.

(28) Wenn Peter da gewesen wäre, wäre es eine nette Party geworden.
$^{?}$Also: Wenn Peter da gewesen und vom Balkon gestürzt wäre, dann wäre es eine nette Party geworden.

(29) Wenn Peter in Frankreich leben würde, würde er nicht in Paris wohnen.
$^{?}$Also: Wenn Peter in Paris wohnen würde, würde er nicht in Frankreich leben.

Wir können uns leicht Situationen vorstellen, in denen die Prämissen jeweils wahr sind. Mindestens einige dieser Situationen sind aber solche, in denen die jeweilige Konklusion falsch ist. Die Folgerungsmuster, die durch (27), (28) bzw. (29) instantiiert werden, nämlich (Trans), (Verst) und (Kontrap), sind daher nicht gültig für kontrafaktische Konditionale. Diese Muster sind aber gültig für jede strikte Implikation. Also können kontrafaktische Konditionale keine strikten Implikationen sein.

Strikte und variabel strikte Konditionale. Als Teil unseres Vorschlags, kontrafaktische im Sinne strikter Konditionale zu deuten, hatten wir eine bestimmte Interpretation der Art von Notwendigkeit, die bei kontrafaktischen Konditionalen im Spiel sein könnte, unterstellt: Im strikten Konditional $\Box(A \to B)$ sollte der Operator $\Box$ zur Betrachtung der materialen Implikation $A \to B$ in "plausiblen" Welten auffordern. D.h. wenn wir fragen, ob $A \sqsupset B$ in einer Welt a wahr ist, dann sollten wir – so lautet der Vorschlag – nach der Wahrheit von $A \to B$ nicht in schlechthin allen Welten fragen, sondern nur in solchen, die aus der Sicht von a als *plausible* alternative Möglichkeiten in Frage kommen. Aber ganz unabhängig davon, daß die kontrafaktischen Konditionale aus den soeben dargelegten Gründen gar keine strikten Konditionale *irgendeiner* Art sein können, kann dieser Vorschlag auch aus einem anderen Grunde nicht richtig sein. Denn wenn das Antezedens A selbst eine Annahme macht, die wir in a für unplausibel halten, dann wäre dem Vorschlag zufolge, A in allen zu betrachtenden, d.h. a-plausiblen Welten falsch und also $A \to B$ in allen diesen Welten wahr; also wäre $A \sqsupset B$ in a wahr, gleichgültig wofür B steht.

An einem Beispiel sei das illustriert. Zu dem Satz

(30) Wenn Peter über die Brüstung steigen würde, dann würde er sich in Lebensgefahr begeben

können wir uns leicht eine Situation denken, in der er etwas Wahres aussagt. Im Sinne der oben betrachteten Hypothese könnte man etwa sagen: In allen plausiblen alternativen Möglichkeiten ist das einfache Konditional (material verstanden)

(31) Wenn Peter über die Brüstung steigt, dann begibt er sich in Lebensgefahr

wahr. Unplausibel – und daher von der Betrachtung auszuschließen – sind zum Beispiel Welten, in denen Peter wie eine Feder schweben kann. Aber was ist dann mit Sätzen wie

(32) Wenn Peter über die Brüstung steigen würde und schweben könnte, dann würde er sich in Lebensgefahr begeben ?

Die Hypothese forderte, daß wir mit jeder Welt a eine Menge $R(a) = \{b \in W : Rab\}$ von plausiblen Alternativen verbinden. In diesem fixen Bereich $R(a)$ müssen sich nun die materialen Konditionale bewähren, so der Vorschlag. Damit der Satz (30) wahr ist (was er ja intuitiv ist), mußten wir annehmen, daß $R(a)$ nur Welten enthält, in denen Peter den Gesetzen der Schwerkraft unterworfen ist, also nicht schweben kann. Wenn wir jetzt mit dieser Annahme über $R(a)$ die Wahrheit von (32) beurteilen möchten, dann ist im Bereich $R(a)$ das Antezedens des materialen Konditionals

(33) Wenn Peter über die Brüstung steigt und schweben kann, dann begibt er sich in Lebensgefahr

durchweg falsch und damit (33) genauso wahr wie

(34) Wenn Peter über die Brüstung steigt und schweben kann, dann begibt er sich nicht in Lebensgefahr.

Aber nur (34) scheint ein wahres kontrafaktisches Konditional zu entsprechen, nämlich: Wenn Peter über die Brüstung steigen würde und schweben könnte, dann würde er sich nicht in Lebensgefahr begeben.

Das Problem liegt offenbar darin, daß strikte Konditionale uns nicht erlauben, die zu betrachtenden Möglichkeiten *nach Maßgabe des Antezedens* zu variieren. Wenn wir uns fragen, ob es möglicherweise gefährlich ist, wenn Peter über die Brüstung steigt, dann fassen wir einen anderen Bereich von Möglichkeiten ins Auge, als wenn wir uns fragen, ob das möglicherweise für einen schwebenden Peter gefährlich ist. Kontrafaktische Konditonale stellen uns vor die Aufgabe, uns das Antezedens in naheliegender Weise vorzustellen. Was wir uns dazu vorstellen, d.h. welche Möglichkeiten in Betracht zu ziehen sind, hängt aber nicht nur davon ab, wie unsere Welt beschaffen ist, sondern auch davon, welche Modifikationen der Welt das Antezedens verlangt. Mit anderen Worten, wenn wir die Wahrheit eines kontrafaktischen Konditionals $A \sqsupset B$ in a beurteilen wollen, dann hängt der Bereich der plausiblen Alternativen nicht allein von a ab, sondern auch von den Welten, die durch das Antezedens A festgelegt werden. Statt mit $R(a)$ sollten wir also den zu betrachtenden Bereich besser mit $R([\![A]\!], a)$ angeben, wobei $[\![A]\!] = \{b : b \models A\}$ die Menge der A-verifizierenden Welten im jeweils betrachteten Modell ist.

Statt $b \in R([\![A]\!], a)$ wollen wir $R_{[\![A]\!]}ab$ schreiben. Unsere Überlegungen münden in diese Wahrheitsbedingung für kontrafaktische Konditionale:

$$(*) \qquad a \models A \sqsupset B \text{ gdw } \forall b: \text{ wenn } R_{[\![A]\!]}ab \text{ und } b \models A, \text{ dann } b \models B.$$

Ein kontrafaktisches Konditional $A \sqsupset B$ ist also genau dann wahr, wenn das materiale Konditional $A \to B$ durchweg ("strikt") wahr ist in einem Bereich alternativer Möglichkeiten, der mit dem Antezedens variiert. Deshalb werden kontrafaktische Konditionale nach diesem Vorschlag auch *variabel strikte Konditionale* genannt.

Wie soll man sich den Bereich von Möglichkeiten, in denen das Antezedens eines Konditionals auf naheliegende Weise verwirklicht ist, vorstellen? Zumindest sicher nicht anders als so, daß in diesen Welten das Antezedens wahr sein muß, also:

$$(\text{id}) \qquad \text{Wenn } R_{[\![A]\!]}ab, \text{ dann } b \in [\![A]\!].$$

Unter dieser Minimalbedingung können wir $(*)$ vereinfachen zu:

$$(\sqsupset) \qquad a \models A \sqsupset B \text{ gdw } \forall b: \text{ wenn } R_{[\![A]\!]}ab, \text{ dann } b \models B.$$

Nur unter der Annahme, daß Rahmen die Bedingung (id) erfüllen, ist der Übergang von $(*)$ zu $(\sqsupset)$ inhaltlich sinnvoll. Dennoch wollen wir im nächsten Abschnitt zunächst die Klasse *aller* Rahmen, also auch solcher, die (id) nicht erfüllen, betrachten und in den Modellen auf solchen Rahmen Konditionale mittels $(\sqsupset)$ interpretieren. Das hat den Vorteil, daß wir so Resultate hinsichtlich eines Basissystems erbringen können, auf denen wir auf einfache Weise aufbauen können.

3. Semantik kontrafaktischer Konditionale und die Basislogik CK

Variabel relationale Rahmen. Wir betrachten eine aussagenlogische Sprache mit einem funktional vollständigen Satz wahrheitsfunktionaler Junktoren (z.B. $\neg$ und $\wedge$), die um den kontrafaktischen Junktor $\sqsupset$ erweitert ist. Eine solche Sprache interpretieren wir in *variabel relationalen Rahmen* der Art

$$(W, R),$$

wobei R nun eine *drei*stellige Relation $R \subseteq \wp(W) \times W \times W$ ist. Für $X \in \wp(W)$ und $a, b, \in W$ schreiben wir statt $(X, a, b) \in R$ kurz: $R_X ab$. Modelle (W, R, I) induzieren wieder eine Wahrmacherrelation $\models_I$ (den Index lassen wir meist weg) mit dieser Klausel für $\sqsupset$-Formeln:

$$(\sqsupset) \qquad a \models A \sqsupset B \text{ gdw } \forall b: \text{ wenn } R_{[\![A]\!]}ab, \text{ dann } b \models B.$$

Der Ausdruck $[\![A]\!]$ (eigentlich $[\![A]\!]_I$) steht für das, was man in der Semantik möglicher Welten eine *Proposition* nennt. Eine Proposition ist eine Menge von Welten – in unserem Fall:

$$[\![A]\!]_I = \{a \in W : a \models_I A\},$$

also die Menge aller Welten, in denen die Formel A unter der Interpretation I wahr ist. Wahrheit im Modell, Gültigkeit und Folgerung sind wie auf p. 206 für die Modellierung der strikten Implikation definiert.

> *Anmerkung zu allgemeinen Rahmen.* Ist jede Menge von Welten eine Proposition? Wir werden hier annehmen, daß das so ist. Aber diese recht starke Annahme ist eigentlich nicht nötig. Es würde genügen, wenn die Struktur der Propositionen der Struktur der Sprache entspricht. Um dies zu garantieren, müßten wir die Definition eines Rahmens erweitern. In einem *allgemeinen Rahmen* (*general frame*), (W, Π, R), ist Π, die Menge der Propositionen, eine Teilmenge von $\wp(W)$, welche unter bestimmten Mengenoperationen – den Operatoren der Sprache entsprechend – abgeschlossen ist. Für manche Zwecke ist es besser, von allgemeinen Rahmen auszugehen; siehe [31, 50] und eine Anwendung in der Konditionallogik in [263]. Wir werden allgemeine Rahmen hier nicht weiter behandeln.

In der Art semantischer Theorie, die wir hier zugrundelegen, repräsentiert die Menge $[\![A]\!]$ die Bedeutung der Formel A (in einem Modell). Anders gesagt: Die Bedeutung eines Satzes ist die Proposition, die er ausdrückt. Natürlich können zwei syntaktisch verschiedene Sätze A und A' in genau denselben Welten wahr sein. Dann ist $[\![A]\!] = [\![A']\!]$, d.h. die beiden Sätze drücken dieselbe Proposition aus.

Obwohl in einem Ausdruck wie $R_{[\![A]\!]_I}ab$ auf eine Interpretation I der Formel A Bezug genommen wird, ist die Relation R doch vollständig Teil eines sprachunabhängigen Rahmens. Die erste Koordinate des Tripels $([\![A]\!]_I, a, b)$ wird hier zwar mithilfe einer Interpretation I festgelegt, aber das ist nicht wesentlich für die Definition der Relation R; wesentlich ist nur, daß die erste Koodinate aus $\wp(W)$ stammt. Deshalb können wir allgemeine Bedingungen für R allein mit Bezug auf Rahmen, nicht auf Interpretationen, d.h. Modellen formulieren.

Manche Autoren (z.B. Stalnaker [275] und Chellas [52, 53]) ziehen die Darstellung der Rahmen mithilfe einer *Auswahlfunktion* vor. Rahmen sind dann von der Art (W, f) mit

$$f : \wp(W) \times W \to \wp(W)$$

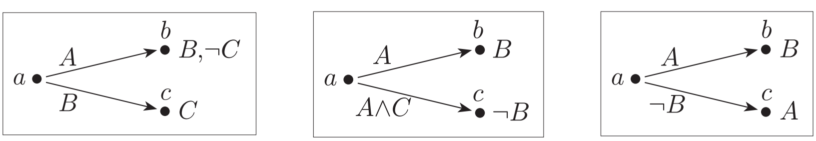
und in den Modellen ist die Bedingung ($\sqsupset$) ersetzt durch

$$(\sqsupset') \qquad a \models A \sqsupset B \text{ gdw } \forall b: \text{ wenn } b \in f([\![A]\!], a), \text{ dann } b \models B.$$

Die Funktion f wählt die plausiblen Antezedens-Welten aus. Das ist lediglich eine Notationsvariante unserer Definitionen, denn wir können von der einen zur jeweils anderen Darstellung leicht übergehen aufgrund der (definitorischen) Äquivalenz

$$(X, a, b) \in R \text{ gdw } b \in f(X, a).$$

Bevor wir uns der Frage zuwenden, wie sich die logisch gültigen Sätze axiomatisch darstellen lassen, wollen wir uns zunächst vergewissern, daß die Wahrheitsbedingung ($\sqsupset$) besser auf kontrafaktische Konditionale paßt als die Bedingung ($\twoheadrightarrow$) für strikte Konditionale. Wir fragen hier insbesondere danach, ob die Sequenzen (Trans), (Verst) oder (Kontrap) (s. Beobachtung 1) sich auch nach dem jetzt vorliegenden Vorschlag als gültig erweisen. Dazu betrachten wir die folgenden, diagrammatisch skizzierten Modelle:



Diagr. 1 2 3

Ein mit einer Formel A beschrifteter Pfeil soll die $R_{[\![A]\!]}$-Relation darstellen. Nach der oben besprochenen Minimalbedingung ist daher A wahr an jedem Punkt auf den ein A-Pfeil weist. (Beispielsweise ist im Diagramm 1 A am Punkt b und B am Punkt c wahr.)

Diagramm 2 zeigt ein Gegenbeispiel zu (Verst), $A \sqsupset B \models A \wedge C \sqsupset B$. Hier sei b der einzige Punkt, der in der $R_{[\![A]\!]}$-Relation zu a steht (angedeutet durch den oberen Pfeil) und B sei an diesem Punkte wahr. Dann $a \models A \sqsupset B$. Der Punkt c sei der einzige Punkt, der zu a in der $R_{[\![A\wedge C]\!]}$-Relation steht (angedeutet durch den unteren Pfeil) und an diesem Punkt möge B falsch sein. Dann gilt nicht $a \models A \wedge C \sqsupset B$.

Das Gegenbeispiel wäre nicht möglich, wenn wir forderten, daß wenn $R_{[\![A\wedge C]\!]}ab$, dann $R_{[\![A]\!]}ab$. Aber genau das sollten wir nicht tun. Denn eine Welt, die unter der Annahme $A \wedge C$ naheliegend ist, ist es unter der Annahme A möglicherweise nicht. So ist, in unserem Beispiel, ein mißlungener Abend ein naheliegendes Szenario unter der Annahme, daß Peter kommt und vom Balkon stürzt. Aber allein Peters Anwesenheit garantiert keinen mißlungenen Abend.

Da $[\![A \wedge C]\!] = [\![A]\!] \cap [\![C]\!]$, können wir unsere Beobachtung etwas verallgemeinern und festhalten, daß (Verst) ungültig bleibt, solange wir nicht fordern, daß

$$R_{X\cap Y} \subseteq R_Y.$$

Da $X \cap Y \subseteq Y$, so repräsentiert $X \cap Y$ eine spezifischere, d.h. logisch stärkere Proposition als Y. (Daß Peter zur Party kommt, läßt mehr Möglichkeiten zu, als daß Peter kommt und stürzt.) Unser Gegenbeispiel richtet sich also allgemeiner gegen das Prinzip,

$$B \sqsupset C \models A \sqsupset C, \text{ falls } A \models B,$$

und damit gegen diese Bedingung:

(mono) $$R_X \subseteq R_Y, \text{ falls } X \subseteq Y.$$

Diagramm 1 zeigt ein Gegenbeispiel zu (Trans), $A \sqsupset B, B \sqsupset C \models A \sqsupset C$. Wir nehmen an, daß $R_{[\![A]\!]}ab$ mit $b \models B$ und $R_{[\![B]\!]}ac$ mit $c \models C$. Wenn, wie im Bild angedeutet, b und c die einzigen Punkte sind, die in der jeweiligen Relation zu a stehen, dann folgt aus diesen Annahmen, daß $a \models A \sqsupset B$ und $a \models B \sqsupset C$. Aber am Punkt b ist A wahr und C falsch. Also nicht $a \models A \sqsupset C$.

Schließlich zeigt Diagramm 3 ein Gegenbeispiel zu (Kontrap), $A \sqsupset B \models \neg B \sqsupset \neg A$. Der obere Teil des Diagramms verifiziert $a \models A \sqsupset B$, der untere falsifiziert $a \models \neg B \sqsupset \neg A$. (Der Leser möge nun die intuitiven Gegenbeispiele (27-29) in Beziehung setzen zu den gerade besprochenen formalen Gegenbeispielen. Ferner möge man sich Bedingungen für die dreistellige Relation R (bzw. die Funktion f) überlegen, welche die Folgerungen (Trans) und (Kontrap) garantieren könnten.)

Die Basislogik CK. Die kleinste hier betrachtete Logik kontrafaktischer Konditionale basiert auf den klassischen Tautologien (im vollen Vokabular, d.h. inklusive Formeln, in denen der Junktor $\sqsupset$ vorkommt). Dem wird ein Axiomenschema sowie eine schematische Regel für Konditionale hinzugefügt. Die Logik ist unter Modus Ponens abgeschlossen:

τ $\quad$ Alle Tautologien

K. $$A \sqsupset (B \to C) \to ((A \sqsupset B) \to (A \sqsupset C))$$

RN. $$\frac{B}{A \sqsupset B}$$

MP. $$\frac{A, A \to B}{B}$$

Die Bezeichnungen K und RN sollen an die gleich benannten Schemata der Basislogik **K** für den Notwendigkeitsoperator $\Box$ erinnern:[5]

K. $$\Box(B \to C) \to \Box B \to \Box C$$

RN. $$\frac{B}{\Box B}$$

Man beachte, daß in den konditionallogischen Schemata der Formelteil $A \sqsupset$ die Stelle von $\Box$ in den modallogischen Schemata eingenommen hat. Tatsächlich hat ja auch schon die semantische Analyse gezeigt, daß $A \sqsupset B$ in etwa bedeutet: B ist notwendig relativ zu A. D.h. jeder Formelteil $A \sqsupset$ verhält sich auch semantisch wie ein Notwendigkeitsoperator. Das wird besonders augenfällig, wenn wir für einen Moment die Notation $A \sqsupset B$ durch $[A]B$ ersetzen (wie in [52] erstmals vorgeschlagen). Jetzt sieht die Sache so aus: Während wir in der (einfachen) Modallogik immer nur *einen* Notwendigkeitsoperator $\Box$ betrachten, so haben wir es in der Konditionallogik mit so *vielen* Notwendigkeitsoperatoren $[A]$ zu tun, wie es Formeln A gibt. Die strenge syntaktische und semantische Analogie zwischen $[A]B$ und $\Box B$ gibt den Schlüssel für die folgenden Resultate in die Hand.

[5] Zur Erinnerung (siehe das Kapitel III.5 über Modallogik): Wenn wir τ und MP um die Schemata K und RN für $\Box$ erweitern, erhalten wir eine Axiomatisierung der Menge aller Formeln, die in beliebigen Kripke-Modellen wahr sind. Das ist das System **K**.

Bevor wir diese Resultate vorstellen, sehen wir uns noch eine alternative Axiomatisierung von **CK** an. Auch diese basiert auf den Tautologien und ist unter Modus Ponens abgeschlossen. Hinzu kommen zwei Regeln für Konditionale:

REA. $$\frac{A \leftrightarrow A'}{(A \sqsupset B) \to (A' \sqsupset B)}$$

RK. $$\frac{B_1 \wedge \cdots \wedge B_n \to B}{(A \sqsupset B_1) \wedge \cdots \wedge (A \sqsupset B_n) \to (A \sqsupset B)}\ n \geq 0.$$

Nach dem soeben Gesagten, ist die Regel RK unschwer zu erkennen als konditionallogische Version der bekannten modallogischen Regel. Die Regel REA (engl. *replacement of equivalent antecedents*) ist dagegen anderer Art: In ihr wird nicht dasselbe Antezedens Formeln vorangestellt, sondern es wird auf verschiedene Antezedens-Formeln Bezug genommen. REA fordert, daß nur die logische Kraft eines Antezedens zählen soll: Die Art und Weise, wie diese logische Kraft ausgedrückt wird, soll für die Wahrheit eines Konditionals keine Rolle spielen – logisch gleichwertige Beschreibungen sind im Wenn-Teil eines Konditionals austauschbar. (Man überzeuge sich davon, daß die zwei vorgestellten Axiomatisierungen äquivalent sind.)

Der Nachweis der *Richtigkeit* von **CK** in variabel relationalen Rahmen ist eine einfache Übung. Für den Beweis der *Vollständigkeit* bedienen wir uns der Methode kanonischer Modelle. Wir nehmen an, eine Formel A sei *kein* Theorem von **CK** und konstruieren ein (kanonisches) variabel relationales Modell, in dem A an einem Punkt falsch wird. Die Methode wurde im Kapitel III.5 über Modallogik im Detail geschildert. Wir werden uns hier kürzer fassen und unser Augenmerk auf die wesentlich konditionallogischen Momente richten.

Eine Menge X von Formeln ist genau dann *maximal konsistent* im Sinne von **CK**, wenn sich (a) in **CK** aus X kein Widerspruch, d.h. keine Formel der Form $A \wedge \neg A$ ableiten läßt, und (b) aus jeder Erweiterung von X sich ein solcher Widerspruch in **CK** ableiten läßt.

Wir wissen nach Lindenbaums Lemma (siehe das Einleitungskapitel), daß für jede Ableitungsrelation $\vdash$ gilt:

> Jede (unter $\vdash$) konsistente Formelmenge läßt sich zu einer (unter $\vdash$) maximal konsistenten Menge erweitern.

Das Lindenbaum-Lemma gilt also auch für die Ableitungsrelation von **CK**. Im Rest dieses Abschnitts meinen wir mit der (maximalen) Konsistenz einer Formelmenge immer diese Eigenschaft im Sinne von **CK**.

Wir wollen nun eine Abkürzung vereinbaren (die *A-Entkonditionalisierung* einer Formelmenge X):

$$X^A = \{B : A \sqsupset B \in X\}.$$

Der Grund für die Einführung dieser Abkürzung ist folgender. Im nächsten Lemma wollen wir zeigen, daß sich maximal konsistente Mengen im wesentlichen wie die Punkte eines Modells verhalten. Dazu brauchen wir insbesondere auch die Bedingung 4 (im Lemma unten), deren Sinn sich aus dieser Definition einer Relation R ergibt:

Def. $R^{\mathbf{CK}}$ $\qquad R_{[A]}XY$ gdw $X^A \subseteq Y$,

wobei $[A]$ für die Menge aller maximal konistenten Mengen, welche A enthalten, steht.

LEMMA 2. *Es seien X und Y maximal konsistente Formelmengen. Dann gilt für alle Formeln A und B:*

1. $A \in X$ gdw $X \vdash A$;
2. $\neg A \in X$ gdw $A \notin X$;
3. $A \wedge B \in X$ gdw $A \in X$ und $B \in X$;
4. $A \sqsupset B \in X$ gdw für alle Y: $X^A \subseteq Y \Rightarrow B \in Y$.

BEWEIS. Für die Beweise der Behauptungen 1–3. sei der Leser auf die Verifikation der M-Bedingungen im Kapitel III.5 über Modallogik verwiesen. Nach Def. $R^{\mathbf{CK}}$ ist 4 gleichbedeutend mit

$A \sqsupset B \in X$ gdw für alle Y: $R_{[A]}XY \Rightarrow B \in Y$.

Die Behauptung wird ganz analog wie im Fall der Modallogik bewiesen:

LR: Wenn $A \sqsupset B \in X$, dann ist $B \in X^A$, welche Menge in Y enthalten ist.

RL: Wir nehmen an (1) $B \in Y$ für alle maximal konsistenten $Y \supseteq X^A$ und, für *reductio*, (2) $A \sqsupset B \notin X$. Es sei nun $X' = X^A \cup \{\neg B\}$. X' ist konsistent. (Denn anderenfalls gäbe es $B_1, ..., B_n \in X^A$ so, daß $\vdash B_1 \wedge \cdots \wedge B_n \to B$. Daraus würde nach RK folgen $\vdash (A \sqsupset B_1) \wedge \cdots \wedge (A \sqsupset B_n) \to (A \sqsupset B)$. Und da $(A \sqsupset B_1), ..., (A \sqsupset B_n) \in X$, hätten wir $A \sqsupset B \in X$ (Behauptung 1 des Lemmas) – entgegen unserer Annahme (2).) Nach Lindenbaums Lemma können wir X' zu einer maximal konsistenten Menge X^* erweitern. Da $\neg B \in X' \subseteq X^*$, so haben wir $B \notin X^*$ (Behauptung 2 des Lemmas). Da aber $X^A \subseteq X' \subseteq X^*$, so ist X^* eine maximal konsistente Erweiterung von X^A, welche, nach (1), B enthält – Widerspruch.

■

Lemma 3. *Das kanonische Modell für* **CK**,

$$\mathcal{M}^{\mathbf{CK}} = (W^{\mathbf{CK}}, R^{\mathbf{CK}}, I^{\mathbf{CK}}) \text{ mit}$$

(a) $W^{\mathbf{CK}} = \{X \subseteq \text{FML} : X \textit{ ist maximal konsistent}\}$,

(b) $R^{\mathbf{CK}}_{[A]}ab$ *gdw* $a^A \subseteq b$, *und*

(c) $I^{\mathbf{CK}}(P, a) = 1$ *gdw* $P \in a$ ($\forall P \in ATM$),

ist ein Modell im hier zu betrachtenden Sinne, d.h. $(W^{\mathbf{CK}}, R^{\mathbf{CK}})$ *ist ein variabel relationaler Rahmen und* $I^{\mathbf{CK}}$ *ist eine Interpretation auf einem solchen Rahmen.*

Beweis. Die in (a) definierte Menge ist offensichtlich nicht leer und auch die Definition (c) ist unproblematisch.

In der Definition (b) darf die Relation $R_{[A]}$ (Superskript fortgelassen) nicht davon abhängen, wie wir die Indexmenge [A] syntaktisch repräsentieren; d.h., wenn $[A] = [A']$, dann brauchen wir eine Garantie, daß $R_{[A]} = R_{[A']}$. Nun sieht die rechte Seite der Definition (b) ausgeschrieben so aus:

$$A \sqsupset B \in a \;\Rightarrow\; B \in b. \tag{$*$}$$

Die gewünschte Garantie erhalten wir, wenn wir aus $[A] = [A']$ auf

$$A' \sqsupset B \in a \;\Rightarrow\; B \in b \tag{$\dagger$}$$

schließen dürfen (und umgekehrt). Den Schluß ermöglicht die Regel REA. Denn wenn $[A] = [A']$ im kanonischen Modell, dann ist $A \leftrightarrow A'$ in allen Punkten des kanonischen Modells enthalten. Der Schnitt aller Punkte des kanonischen Modells ist aber genau **CK**, d.h. $A \leftrightarrow A'$ ist ein Theorem von **CK**. Also ist, nach REA, auch

$$(A \sqsupset B) \leftrightarrow (A' \sqsupset B)$$

ein Theorem von **CK**. Es folgt, daß $A \sqsupset B \in a$ gdw $A' \sqsupset B \in a$ (für beliebige Punkte a des kanonischen Modells). Also sind ($*$) und ($\dagger$) äquivalent. ■

Lemma 4. *Sei* (W, R, I) *das kanonische Modell für* **CK**. *Es sei eine Relation* $\models \subseteq W \times \text{FML}$ *so definiert:*

$$a \models A \textit{ gdw } A \in a.$$

Dann ist die Relation $\models$ eine Erfüllungsrelation.

Beweis. Induktion über den Aufbau einer Formel A.
Basis $A = P$. Nach Definition: $P \in a$ gdw $I(P, a) = 1$ gdw $a \models P$.
Fall $A = \neg B$. Siehe Lemma 2.2.
Fall $A = B \wedge C$. Siehe Lemma 2.3.
Fall $A = B \sqsupset C$. Zz ist $a \models B \sqsupset C$ gdw $\forall b : R_{[B]}ab \Rightarrow b \models C$, d.h.

$$B \sqsupset C \in a \text{ gdw } \forall b : R_{[B]}ab \Rightarrow C \in b,$$

wobei a und b maximal konsistent sind. Nach der Def. des kanonischen Modells ist das die Äquivalenz

$$B \sqsupset C \in a \text{ gdw } \forall b : a^B \subseteq b \Rightarrow C \in b,$$

welche wir in Lemma 2.4 bewiesen haben. ∎

Das *Argument für die Vollständigkeit* von **CK** im Hinblick auf die Klasse der variabel relationalen Rahmen ist nun recht einfach. Es sei A eine beliebige Formel, die kein Theorem von **CK** ist. Dann ist $X = \mathbf{CK} \cup \neg A$ eine konsistente Menge. Diese läßt sich zu einer maximal konsistenten Menge X^* erweitern (Lindenbaums Lemma). Die Menge X^* ist ein Punkt im kanonischen Modell, an dem A falsch ist (Lemma 2). Also gibt es ein Modell (Lemma 3), welches A falsifiziert (Lemma 4) und somit ist A nicht gültig in der Klasse der variabel relationalen Rahmen. Kontrapositiv ausgedrückt: Wenn A gültig ist in diesen Rahmen, dann ist A ein Theorem von **CK**.

4. Erweiterungen des Basissystems CK

Die Variabilität der Relation R in Abhängigkeit von der Proposition, die das Antezedens eines kontrafaktischen Konditionals repräsentiert, hat, wie wir gesehen haben, die gewünschte Wirkung: Die drei typischen Folgerungsmuster, die für diese Konditionale nicht gelten, lassen sich in variabel relationalen Modellen widerlegen. Für die Basislogik **CK** sind die entsprechenden Ableitbarkeitsbehauptungen daher falsch.

Damit es sich bei der Relation R_X um kein bloß technisches Mittel handelt, die gewünschten Resultate herbeizuführen, sollten wir in der Lage sein, eine Interpretation der Relation anzugeben, die es erlaubt, aus den auf p. 211 gezeigten, abstrakten Gegenmodellen zu (Trans), (Verst) und (Kontrap), konkrete und intuitiv überzeugende Gegenbeispiele zu gewinnen. Die Semantik würde dann *erklären*, wie es zu diesen Gegenbeispielen kommen kann.

Wir haben eine solche Erklärung bereits angedeutet. Wenn wir die Wahrheit eines Konditionals $A \sqsupset B$ in einer Welt a beurteilen wollen, dann müssen wir Welten betrachten, in denen das Antezedens A wahr ist und uns fragen, ob in diesen Welten auch das Konsequens B wahr ist. Welche Welten sind zu betrachten? Sicherlich sind das nicht *alle* A-Welten, denn dann wäre das kontrafaktische Konditional ein striktes. Wenn wir eine Annahme machen, von der wir glauben, daß sie kontrafaktisch sei, dann fassen wir vielmehr Welten ins Auge, die sich von der aktualen Welt nicht allzu sehr unterscheiden sollen. Die Annahme soll in diesen Welten natürlich wahr sein, aber ansonsten soll alles möglichst so sein, wie es in der aktualen Welt ist. Das heißt, wir betrachten Welten, die der aktualen Welt so ähnlich wie möglich sind unter der Bedingung, daß in ihnen das Antezedens wahr ist. $R_{[\![A]\!]}$ sollte also von der Welt a zu allen Welten b führen, in denen A wahr ist und ansonsten a so ähnlich ("nahe") wie möglich sind. In diesen Welten schauen wir nach, ob auch das Konsequens B wahr ist.[6]

Wir können und sollten uns nun fragen, welche Bedingungen R erfüllen muß, damit die Relation so, wie gerade vorgeschlagen, interpretiert werden kann. Bisher haben wir nichts weiter über R angenommen. Die Theorie dieser allgemeinsten Klasse variabel relationaler Rahmen ist die Basislogik **CK**. Wir betrachten nun naheliegende Bedingungen für R und zeigen, welche Axiomenschemata durch diese Bedingungen logisch gültig werden. Es wird in den betrachteten Fällen auch das Umgekehrte gelten (mit einer Ausnahme): Die Axiomenschemata werden nur dann gültig, wenn die Rahmen die entsprechenden Bedingungen erfüllen. "Entsprechung" hat hier den präzisen Sinn einer *Korrespondenz* so, wie wir diesen Begriff im Kapitel III.8 über Modallogik eingeführt haben. Der Nachweis von Korrespondenzen in der Konditionallogik ist nicht anders als in der Modallogik. Details werden wir daher hier nicht erörtern. Als Bezeichnungen für die Schemata und korrespondierenden Bedingungen dienen die in der Literatur üblichen.

---

[6] Ähnlich – in welcher Hinsicht? Viele Konditionale können wir beurteilen ohne über die relevanten Hinsichten der Ähnlichkeit genauer Rechenschaft ablegen zu müssen. Aber wie steht es mit dem folgenden Konditional?:

> Hätte Breschnjew auf den roten Knopf gedrückt, dann sähe die Welt heute ganz anders aus.

Einerseits sind – zumindest auf den ersten Blick – Welten, in denen Breschnjew auf den Knopf drückt und dann der Auslösemechanismus versagt, der aktualen Welt viel ähnlicher als solche, in denen es zu einem Nuklearschlag kommt. Andererseits haben wir keinen Grund anzunehmen, der Mechanismus würde, falls gedrückt, nicht funktionieren. So erfordert ein nicht-funktionierender Mechanismus Fehlerquellen und damit Welten, die unserer Welt unähnlicher sind als Welten, in denen der Mechanismus funktioniert. Da wir geneigt sind, das Konditional als wahr zu beurteilen, muß hier ein Sinn von Ähnlichkeit angenommen werden, der der zweiten und nicht der ersten Überlegung den Vorzug gibt. Vgl. dazu Lewis [184].

Identität. Ein kontrafaktisches Konditional lädt ein, anzunehmen, das Antezedens sei wahr. Zumindest diese minimale Bedingung sollte die Relation R daher erfüllen: Sie sollte nur zu Welten führen, in denen das, was das Antezedens beschreibt, auch der Fall ist. Wir können diese Bedingung kurz so notieren (für alle Welten a und b, sowie beliebige $X \subseteq W$):

(id) $$\text{Wenn } R_X ab, \text{ dann } b \in X.$$

In Modellen, welche (id) nicht erfüllen, läßt sich das Identitäts-Schema

ID. $$A \sqsupset A$$

falsifizieren. Umgekehrt ist die Bedingung hinreichend dafür, daß ID an allen Punkten eines Modells wahr wird.

Die Bedingung (id) ist sicher eine offensichtliche Bedingung, die man von Rahmen unter der intendierten Interpretation der Relation R erwartet. Genauso dürfte das Schema ID zum Grundbestand jeder Konditionallogik gehören. Jedoch entfalten im Rahmen der hier betrachteten Semantik selbst so unschuldig scheinende Prinzipien wie (id)/ID unerwartete und deutlich weniger akzeptable Wirkungen. Im hier betrachteten Fall ist es die Konsequenz, daß mit (id) auch das Schema $\bot \sqsupset A$ gültig wird – wenn Unmögliches der Fall wäre, dann würde Beliebiges der Fall sein. Wir kommen unten (p. 226) darauf zurück.

Schwache Zentrierung. Wir wollen die der Referenzwelt a ähnlichsten Welten betrachten, in denen das Antezedens A wahr ist. Keine Welt ist a ähnlicher als a selbst. Wenn daher das Antezedenz schon in a wahr ist, dann ist die Welt a sicher mit in Betracht zu ziehen.

(cw) $$\text{Wenn } a \in X, \text{ dann } R_X aa.$$

Die Bedingung der schwachen Zentrierung entspricht dem Schema des konditionalen Modus Ponens:

CW. $$(A \sqsupset B) \to (A \to B).$$

Ein Konditional mit wahrem Antezedens kann nur wahr sein, falls das Konsequens wahr ist.

Erfolg. Keine Welt kann derart sein, daß es keine ihr ähnlichen Welten geben kann. Ähnliche Welten sollte es immer geben. Die Suche nach Welten, die unserer so ähnlich wie möglich unter einer bestimmten Bedingung sind, kann also nur an der Bedingung scheitern. Die Bedingung könnte einfach *abwegig* sein, d.h. sie könnte derart sein, daß es keine Welten gibt, die a ähnlich sind *und* dabei die Bedingung erfüllen. Etwas "formaler" ausgedrückt: Die Bedingung X ist abwegig, wenn es kein b gibt mit $R_X ab$. Wenn X in diesem Sinne abwegig ist, dann werden X-Welten auch unter keiner anderen Bedingung erreichbar sein, d.h. für beliebige Bedingungen Y gilt, daß es kein b gibt mit $R_Y ab$ und $b \in X$. Der Gedankengang wirbt für diese Rahmenbedingung:

(mod) $\qquad$ Wenn $\neg\exists b : R_X ab$, dann $\forall Y(\neg\exists b : R_Y ab$ und $b \in X)$.

Gelegentlich bietet es sich an, statt $R_X ab$ auch $b \in R_X(a)$ zu schreiben. In dieser Schreibweise sieht (mod) dann so aus:

(mod) $\qquad$ Wenn $R_X(a) = \emptyset$, dann $\forall Y : R_Y(a) \cap X = \emptyset$.

Diese Bedingung hat nun den folgenden Effekt. Angenommen $\neg A$ ist in a abwegig, d.h. es gibt keine a-ähnlichen Welten, in denen $\neg A$ der Fall wäre. Die Proposition $\neg A$ ist also "ziemlich unmöglich" und A "so gut wie notwendig", was durch die Formel $\neg A \to \bot$ ausgedrückt wird. [7] Dann ist A unter *jeder* Antezedens-Bedingung wahr. In einem Schema können wir das so ausdrücken:

MOD. $\qquad$ $(\neg A \sqsupset \bot) \to (B \sqsupset A)$.

Es ist genau dieses Schema, welches durch die Bedingung (mod) gültig wird.

Denn, angenommen

$$a \models \neg A \sqsupset \bot$$

(zz. $a \models B \sqsupset A$); dann

$$R_{[\![\neg A]\!]}(a) \subseteq [\![\bot]\!].$$

Da $[\![\bot]\!] = \emptyset$, folgt, daß $R_{[\![\neg A]\!]}(a) = \emptyset$. Daraus schließen wir nach (mod), daß

$$R_{[\![B]\!]}(a) \cap [\![\neg A]\!] = \emptyset.$$

D.h. $R_{[\![B]\!]}(a) \subseteq [\![A]\!]$, welches die Wahrheitsbedingung für $a \models B \sqsupset A$ ist. ∎

[7] Tatsächlich definiert Lewis [182, pp. 22-24] einen Notwendigkeitsoperator mit Hilfe des Konditionals so $\Box A := A \sqsupset \bot$. Er nennt diesen modalen Operator einen "äußeren", weil er über den Bereich der äußersten Sphäre eines Sphärensystems (s.u.) definiert ist.

Koinzidenz. Die nächste Bedingung sieht zunächst etwas komplizierter aus, drückt aber eine sehr einfache Idee aus: Wenn alle nächsten A-Welten B erfüllen und alle nächsten B-Welten A erfüllen, dann macht es keinen Unterschied, ob die Antezedens-Welten mit A oder B beschrieben werden: Die nächsten A-Welten sind genau die nächsten B-Welten.

(cso) Wenn $\forall b : R_X ab \Rightarrow b \in Y$ und $\forall b : R_Y ab \Rightarrow b \in X$, dann $\forall b : R_X ab$ gdw $R_Y ab$.

Auch das können wir ein wenig kürzer aufschreiben:

(cso) Wenn $R_X(a) \subseteq Y$ und $R_Y(a) \subseteq X$, dann $R_X(a) = R_Y(a)$.

Dieses Koinzidenz-Prinzip entspricht dem Schema

CSO. $(A \sqsupset B) \wedge (B \sqsupset A) \rightarrow ((A \sqsupset C) \rightarrow (B \sqsupset C))$.

(Starke) Zentrierung. Jede Welt a ist sich selbst strikt ähnlicher als irgendeine von a verschiedene Welt b es sein kann. Das gilt auch dann, wenn a und b vieles gemein haben – z.B. daß beide das zu betrachtende Antezedens A erfüllen. Wenn nun a das Antezedens A erfüllt, dann kann die Menge der a ähnlichsten A-Welten nur eine einzige Welt umfassen, nämlich a selbst. Das ist die Bedingung der starken Zentrierung:

(cs) Wenn $a \in X$ und $R_X ab$, dann $a = b$.

In jedem Fall, in dem es überhaupt eine Welt b mit $R_X ab$ gibt (X also nicht abwegig ist), folgt aus der starken Zentrierung die schwache. Zentrierung ist notwendig und hinreichend für die Gültigkeit des Schemas

CS. $A \wedge B \rightarrow (A \sqsupset B)$.

Gegen CS wird eingewandt, daß es für die Wahrheit eines kontrafaktischen Konditionals nicht hinreichend sein kann, daß Antezedens und Konsequens (zufällig) wahr sind. Beispiele, die CS merkwürdig dastehen lassen, sind schnell bei der Hand.

> Peter rechnet mit Regen und nimmt seinen Schirm mit. Tatsächlich scheint die Sonne. Nach CS ist in dieser Situation der Satz "Wenn es sonnig wäre, würde Peter seinen Regenschirm bei sich haben" wahr.

Andererseits aber entspricht das Schema der Bedingung der starken Zentrierung. Und diese Bedingung ist unter der intendierten Interpretation der

Relation R sehr einleuchtend. Daher ist es naheliegend – wenn auch sicher nicht zwingend – vermeintliche Gegenbeispiele zu CS so zu erklären, daß in diesen Fällen kein falsches Konditional, sondern der falsche Gebrauch eines wahren Konditionals vorliegt. [8]

Einzigkeit. Wie negiert man ein kontrafaktisches Konditional?

(35) Wenn Wagner in Frankreich geboren wäre, wäre er Franzose.

Das ist vermutlich wahr. Indem wir (35) negieren, behaupten wir da nicht

(36) Wenn Wagner in Frankreich geboren wäre, wäre er kein Franzose ?

Und sicher folgt umgekehrt aus (36) die Negation von (35). Also scheint die Negation eines Konditionals nichts anderes als die Negation des Konsequens zu sein:

$^{?}$NEG. $$\neg(A \sqsupset B) \leftrightarrow (A \sqsupset \neg B)$$

– ein Schema, welches äquivalent ist zu

CEM. $$(A \sqsupset B) \vee (A \sqsupset \neg B).$$

CEM ist in der Literatur bekannt als *konditionales Tertium Non Datur* und eines der strittigsten Schemata. [9] Der Streit entzündete sich an direkten Gegenbeispielen zu CEM, welche in der semantischen Perspektive an Schärfe gewinnen. Denn CEM erzwingt die besonders starke Rahmenbedingung

(cem) Wenn $R_X ab$ und $R_X ac$, dann $b = c$.

Die Ähnlichkeitsrelation wird damit funktional, d.h. jede Relation R_X ordnet die Welten so an, daß keine zwei Welten einer gegebenen Welt gleich ähnlich sein können: *In puncto* Ähnlichkeit gibt es kein Unentschieden. [10]

[8] Lewis [182, p. 27]: "Unsere grundsätzliche Erwiderung ist nicht, daß das Konditional [mit wahrem Antezedens und Konsequens] wahr [...] ist, sondern daß es wegen seines wahren Antezedens unrichtig und irreführend ist. [Wahr] ist es zwar, aber das ist nicht der Punkt. Die falsche Information, die durch den Gebrauch einer kontrafaktischen Konstruktion mit einem wahren Antezedens [und Konsequens] zum Ausdruck gebracht wird, überdeckt [...] die Wahrheit des Konditionals." (Übersetzung AF).

[9] Siehe die Diskussion zwischen Lewis und Stalnaker in [182] bzw. [273].

[10] Denselben Effekt hat die von Stalnaker bevorzugte Wahrheitsbedingung

$$a \models A \sqsupset B \text{ gdw } f([\![A]\!], a) \models B.$$

Hier ist f eine Auswahlfunktion (wie auf p. 211 beschrieben), welche statt einer Menge naheliegender Welten eine einzige, naheliegendste als Wert ausgibt; siehe [275].

Daß das wenig plausibel ist, sieht man zum Beispiel an Konditionalen, die so beginnen (Beispiel aus [234, §3]):

(37) Wenn Wagner und Verdi Landsleute wären, dann ...

Welche ist diejenige Welt, die genau wie unsere ist, nur daß darin Wagner und Verdi Landsleute sind? Ist das eine Welt, in der Verdi Deutscher ist oder ist es eine, in der Wagner Italiener ist? Keine der beiden Möglichkeiten scheint unserer Welt ähnlicher als die jeweils andere zu sein. Sie sind, auf ihre jeweilige Weise, unserer Welt gleich nahe. Beide sind als naheliegende Möglichkeiten zu berücksichtigen, wenn wir uns fragen, was der Fall wäre, wenn Wagner und Verdi Landsleute wären. Aber dann kann (cem) nicht richtig sein.

Diesem Argument gegen (cem) steht die *prima facie* plausible Hypothese $^?$NEG über die Negation von Konditionalen entgegen. Aber ist diese Hypothese wirklich überzeugend? Was ein Konditional "Wenn A der Fall wäre, dann würde B wahr sein" falsch macht, ist nicht der Umstand, daß B unter der Annahme A falsch *wäre*, sondern daß B unter der Annahme auch falsch sein *könnte*. Das ist sicher nicht dasselbe.

Neben den *würde*-Konditionalen gibt es auch die *könnte*-Konditionale: "Wenn A der Fall wäre, dann könnte B wahr sein." Solche *könnte*-Konditionale wollen wir so notieren: $A > B$. Für *würde*-Konditionale haben wir gefordert, daß das Konsequens in *allen* naheliegenden Antezedens-Welten wahr sein muß. *Könnte*-Konditionale sind offenbar wahr unter der Bedingung, daß das Konsequens in *einigen* naheliegenden Antezedens-Welten wahr ist, d.h.:

(>) $$a \models A > B \text{ gdw } \exists b : R_A ab \text{ und } b \models B.$$

Mit dieser Wahrheitsbedingung geht unmittelbar ein Vorschlag zum richtigen Verständnis der Negation eines *würde*-Konditionals einher:

NEG. $$\neg(A \sqsupset B) \leftrightarrow A > \neg B.$$

Die Wahrheitsbedingung und das damit einhergehende Schema NEG passen nicht zur Bedingung (cem) bzw. $^?$NEG. Denn wenn wir beides verbinden, dann verschwindet der logische Unterschied zwischen *würde*- und *könnte*-Konditionalen. Also sollten wir ($\sqsupset$) und ($>$) nicht unter die Bedingung (cem) stellen. Und wir sollten uns für NEG statt für $^?$NEG als richtiger Analyse der Negation von *würde*-Konditionalen entscheiden. Denn die richtige Negation von (35) ist nicht (36), sondern

(38) Wenn Wagner in Frankreich geboren wäre, dann *könnte* er (dennoch) nicht Franzose sein.

So können wir mit NEG erklären, was $^?$NEG vordergründig plausibel erscheinen ließ. Es ist richtig, daß man ein Konditional verneint, indem man das Konsequens verneint. Das ist der zutreffende Kern der Hypothese $^?$NEG über die Negation von Konditionalen. Nur wird im Falle eines *würde*-Konditionals das Konsequens des enstprechenden *könnte*-Konditionals verneint, und ein *könnte*-Konditional wird verneint durch die Verneinung des Konsequens im entsprechenden *würde*-Konditional. Die richtigen Verneinungsäquivalenzen sind also:

$$\neg(A \sqsupset B) \leftrightarrow (A > \neg B) \quad \text{bzw.} \quad \neg(A > B) \leftrightarrow (A \sqsupset \neg B).$$

Vorsichtige Verstärkung. Gegenbeispiele zum Prinzip der Verstärkung des Antezedens,[11]

$$(*) \qquad (A \sqsupset B) \to (A \wedge C \sqsupset B),$$

scheinen alle derart zu sein, daß die verstärkende Annahme C die Wahrheit von B unterminieren würde. Aber dann müßte Verstärkung um zumindest solche Aussagen möglich sein, die das Konsequens *nicht* unterminieren. Welche Aussagen sind das?

Betrachten wir noch einmal das folgende Gegenbeispiel zu $(*)$:

(28) (a) Wenn Peter da gewesen wäre (A), wäre es eine nette Party geworden (B).
$^?$Also: (c) Wenn Peter da gewesen und vom Balkon gestürzt wäre ($A \wedge C$), dann wäre es eine nette Party geworden.

Das Beispiel kann gegen $(*)$ nichts ausrichten, wenn in mindestens einer der naheliegenden Möglichkeiten, in denen Peter auf der Party ist, er vom Balkon stürzt. Denn dann ist die Möglichkeit des Sturzes in (a) eingerechnet und daher kann (c) nicht falsch sein.

Wer also (a) behauptet, kann (c) nur ablehnen mit der Begründung, daß unter der Annahme A die Möglichkeit eines Balkonsturzes weit hergeholt ist. Wenn wir diese Begründungsmöglichkeit abschneiden, dann *folgt* (c) aus (a). Die einfachste Weise, die genannte Möglichkeit im Beispiel auszuschließen, geht so:

(b) Wenn Peter da gewesen wäre, hätte er vom Balkon stürzen können.

Wer (a) und (b) behauptet, der ist auch zur Behauptung von (c) verpflichtet.

[11] Wird das Antezedens verstärkt, so wird das Konditional geschwächt, weshalb (kurz) Verstärkung (*strengthening*) manchmal auch (kurz) "Abschwächung" (*weakening*) genannt wird.

Unsere Überlegung können wir nun so verallgemeinern und zusammenfassen: Wenn $A \sqsupset B$ wahr ist (also alle naheliegenden A-Welten auch B-Welten sind) und in einigen naheliegenden A-Welten auch C der Fall ist, dann kann die Präsenz von C in A-Welten die Wahrheit von B nicht unterminieren, und also ist dann auch $A \wedge C \sqsupset B$ wahr. Daß A in naheliegenden Welten C nicht ausschließt, wird durch das *könnte*-Konditional $A > C$ wiedergegeben. Wir haben so ein Argument für die logische Wahrheit des Schemas der *vorsichtigen Verstärkung*,

CV. $$(A \sqsupset B) \wedge (A > C) \rightarrow (A \wedge C \sqsupset B).$$

Das Schema korrespondiert mit der Rahmenbedingung

(cv) Wenn $\exists b : R_X ab$ und $b \in Y$ und $R_{X \cap Y} ac$, dann $R_X ac$.

Das Prinzip der vorsichtigen Verstärkung wird zuweilen skeptisch gesehen. Das folgende Gegenbeispiel – wenn es denn eines ist – stammt von Stalnaker [274]. Brahms, Verdi und Wagner sind keine Landsleute, jedenfalls nicht alle drei. Aber wenn Brahms und Verdi Landsleute wären, dann würden entweder beide Italiener oder beide Deutsche sein. Diese zweite Möglichkeit, wie Brahms und Verdi Landsleute sind, ist eine, in der auch Verdi und Wagner Landsleute sind. Also *könnte* es sein, daß auch Verdi und Wagner Landsleute sind ($v \sim w$), wenn Brahms und Verdi es wären ($b \sim v$):

(39) $A > C : \ b \sim v > v \sim w.$

Andererseits bleibt Wagner Deutscher in allen Welten, die unserer so ähnlich wie möglich sind, bis auf den Umstand, daß Brahms und Verdi Landsleute sind. Die kontrafaktische Annahme, daß Brahms und Verdi Landsleute sind, erzwingt keinen Wechsel der Nationalität Wagners. D.h., wenn Brahms und Verdi Landsleute wären, dann *würde* Wagner (immer noch) Deutscher sein (Dw):

(40) $A \sqsupset B : \ b \sim v \sqsupset Dw.$

Aus (39) $\wedge$ (40) können wir jetzt mit CV schließen auf

(41) $A \wedge C \sqsupset B : \ b \sim v \wedge v \sim w \sqsupset Dw,$

d.h. wenn alle drei Landsleute wären, dann *würde* Wagner Deutscher und also würden alle drei Deutsche sein. Aber das, so möchten wir sagen, ist sicher falsch, denn es gibt mindestens eine andere naheliegende Weise, in der die drei Komponisten Landsleute sein könnten: Sie könnten ja alle Italiener sein.

Haben wir die Gültigkeit von CV widerlegt? Die Prämisse (39) behauptet – ID einmal voraussetzend –, daß es naheliegende Welten gibt, in denen alle drei Komponisten Landsleute sind. Das können sie in diesen Welten auf genau zweierlei Weise sein: Entweder sind sie alle Italiener oder sie sind alle Deutsche. Der erste Fall widerspricht der zweiten Prämisse, (40). Also besagen die beiden Prämissen zusammengenommen, daß die drei nur auf eine Weise Landsleute sein können, nämlich indem sie Deutsche sind. Das ist das, was die Konklusion behauptet und wenn diese Konklusion falsch sein sollte, dann können die beiden Prämissen nicht zugleich wahr sein. D.h. wenn die drei tatsächlich auf *zweierlei* Weise Landsleute sein können (die Negation von (41)), dann schließen diese Möglichkeiten einander aus:

(42) Brahms und Verdi können nur als Deutsche Landsleute sein, und
(43) die drei könnten Landsleute sein.

Der Schluß von (39) und (40) auf (41) ist also völlig in Ordnung, wenn man nur richtig versteht, welche Situation die Prämissen zusammen beschreiben.

Unmögliches Antezedens. Wie gehen wir um mit Konditionalen der Form

IMP $\quad \bot \sqsupset A \quad ?$

In der Basislogik **CK** ist IMP kein Theorem. (Man betrachte zwei Punkte, a und b, mit $R_{[\![\bot]\!]}ab$ und $b \not\models A$. Dann ist $\bot \sqsupset A$ am Punkt a falsch.) Sobald jedoch die Bedingung (id) gilt, muß die Menge $\{b : R_X ab\}$ in X enthalten sein. Da $[\![\bot]\!] = \emptyset$, bedeutet das, daß $\{b : R_{[\![\bot]\!]}ab\} \subseteq \emptyset$. Also gibt es keinen Punkt b so, daß $R_{[\![\bot]\!]}ab$. Damit ist die Wahrheitsbedingung für $a \models \bot \sqsupset A$, nämlich $\forall b : R_{[\![\bot]\!]}ab \Rightarrow b \models A$, an jedem Punkt a auf leere Weise erfüllt und also ist IMP ein gültiges Schema.

Nun scheint es aber so zu sein, daß einige kontrafaktische Konditionale mit unmöglichem Antezedens falsch und andere auf nicht-leere Weise wahr sind – z.B.:

(44) Wenn das System **CK** die richtige Konditionallogik wäre, dann wäre ID ein gültiges Schema. (Falsch.)
(45) Wenn Peter die Quadratur des Kreises gelingen würde, wäre seine Mathematiklehrerin überrascht. (Wahr.)

Die Logik **CK** kann nicht die richtige Konditionallogik sein; dafür ist sie zu schwach. Aber wenn sie es wäre, dann wäre ID *nicht* gültig. Mit (45) liegt dagegen ein Konditional vor, von dem wir uns mühelos vorstellen können, daß es auf eine gehaltvolle Weise wahr ist, im Gegensatz etwa zu dem Konditional

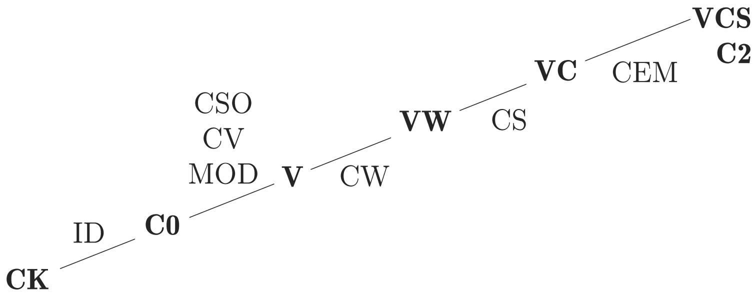
(46) Wenn Peter die Quadratur des Kreises gelingen würde, dann würde er häufiger baden.

Wir können solche Beispiele zum Anlaß nehmen, unsere semantische Theorie zu revidieren. In diesem Fall müßten wir zwischen verschiedenen unmöglichen Welten unterscheiden: Wir müßten unterscheiden zwischen einer (unmöglichen) Welt, in der die Quadratur des Kreises überrascht und einer, in der sie nicht überrascht. Kurzum, wir müßten die Beschränkung auf *mögliche* Welten aufgeben.[12]

Eine andere theoretische Option besteht darin, auf dieses "Paradox der kontrafaktischen Implikation" wie auf die gleichartigen Schwierigkeiten mit der materialen und der strikten Implikation zu reagieren. Wir müßten erklären, warum manche Konditionale, obwohl wahr, nicht behauptbar sind. Von einem Konditional wie (44), zum Beispiel, würden wir sagen, daß es "nicht richtig" sei. Können wir dieses "nicht richtig" im Sinne von "nicht behauptbar obwohl (auf leere Weise) wahr" interpretieren? Warum ist (44) nicht behauptbar? Und warum ist (45) im Gegensatz zu (46) behauptbar, obwohl beide auf bloß leere Weise wahr sind? Eine Antwort müßte hier die einfache semantische Theorie der Wahrheitsbedingungen mit einer komplizierten Theorie der Behauptbarkeitsbedingungen kombinieren. Dagegen scheint unter der zuerst genannten Option die Erklärung recht einfach zu sein: (45) ist behauptbar, weil das Konditional wahr ist, während (44) und (46) nicht behauptbar sind, weil die Konditionale falsch sind. Aber diese einfache Antwort müßten wir uns mit einer erheblichen Komplizierung der semantischen Theorie erkaufen.

Einige bekannte Erweiterungen von CK. Die Logik **CK** ist eine durch ihre einfachen Modelle ausgezeichnete Ausgangsbasis für eine Logik kontrafaktischer Konditionale; sie ist aber sicher selbst keine solche Logik. Es fehlt beispielsweise das Schema ID. Das folgende Diagramm zeigt die Beziehungen zwischen sechs Logiken, die in der Literatur häufig Erwähnung finden. Eine Linie von unten nach oben deutet eine Erweiterung an, wobei die Linie mit den Schemata versehen ist, um welche das unmittelbar darunter stehende System erweitert wird; vgl. auch die Tafel auf p. 240.

[12] Siehe [91] und [42] für einen Einblick in mögliche Strategien für eine Revision der semantischen Theorie.



Während **CK** so etwas wie ein "technischer Anfang" ist, spielt das System **C0** die Rolle einer gesicherten Basis für die Konditionallogik. Auch die Hinzunahme von MOD und CSO ist unkontrovers. Mit der vorsichtigen Verstärkung CV beginnen jedoch die Kontroversen. Noch umstrittener sind schwache und starke Zentrierung, CW und CS, und das konditionale *tertium non datur* CEM. Diese Systeme und ihre charakteristischen Schemata bilden jedoch nur die Eckpunkte für eine Diskussion, die weitaus subtiler und verzweigter ist als es hier dargestellt werden kann.

Das System **VC** ist das von Lewis in [182] bevorzugte System. Es handelt sich dabei um die logische Theorie einer besonders suggestiven Semantik für kontrafaktische Konditionalsätze, der sogenannten (zentrierten) Sphärensysteme.

Übung. Man betrachte die folgenden Schemata, um die man die kleinste normale Modallogik **K** erweitern kann. (Vgl. den Abschnitt 8 über Definierbarkeit im Kapitel III über Modallogik.)

T.	$\Box B \to B$
4.	$\Box B \to \Box\Box B$
B.	$B \to \Box\Diamond B$
5.	$\Diamond\Box B \to \Box B$
D.	$\Box B \to \Diamond B$
4c.	$\Box\Box B \to \Box B$
G.	$\Diamond\Box B \to \Box\Diamond B$

Man ersetze jedes Vorkommen von $\Box$ durch $A \sqsupset$ bzw. von $\Diamond$ durch $A >$. Sind die so entstandenen Schemata plausibel für kontrafaktische Konditionale? Welchen Rahmenbedingungen würden sie entsprechen? Sind einige dieser Schemata in einigen der oben genannten Konditionallogiken enthalten?

5. Lewis' Sphärensemantik

Die Grundidee der hier vorgestellten semantischen Theorie für kontrafaktische Konditionale besteht darin, daß Welten einander mehr oder weniger ähnlich sein können. Um ein Konditional in einer Welt zu beurteilen, sollten wir nur solche Welten betrachten, in denen das Antezedens wahr und alles andere möglichst unverändert ist. Es liegt nahe, sich die Ähnlichkeit bildhaft vorzustellen, in dem wir das Maß der Ähnlichkeit in ein topographisches Maß der Entfernung übersetzen: Je unähnlicher eine Welt der Ausgangswelt ist, umso weiter ist jene von dieser entfernt. So stellen wir uns schließlich Sphären der Ähnlichkeit, um die Ausgangswelt angeordnet, vor – wie eine Matroschka-Puppe oder die Schichten einer Zwiebel. Diese bildliche Vorstellung steht tatsächlich in einer strikten Analogie zu den bisher betrachteten variabel relationalen Rahmen.

Es sei $\mathfrak{S}$ eine Funktion, die jeder Welt $a \in W$ eine Menge $\mathfrak{S}_a$ von Mengen möglicher Welten zuordnet:

$$\mathfrak{S} : W \longrightarrow \wp(\wp(W)),$$

d.h. $\mathfrak{S}_a \subseteq \wp(W)$. Wir nennen $(W, \mathfrak{S})$ einen *Sphärenrahmen*, wenn jede Menge $\mathfrak{S}_a$ ein Sphärensystem um a ist. Die Elemente S eines Sphärensystems $\mathfrak{S}_a$ heißen *Sphären*. Ein *Sphärensystem* $\mathfrak{S}_a$ *um* a soll die folgenden Bedingungen erfüllen:

(Scs) $\mathfrak{S}_a$ ist auf a *zentriert*: $\{a\} \in \mathfrak{S}_a$.
(Sli) $\mathfrak{S}_a$ ist *verschachtelt*: $\forall S, S' \in \mathfrak{S}_a : S \subseteq S'$ oder $S' \subseteq S$.
(Sla) *Limes-Annahme*: $\forall X \subseteq W$: Wenn $X \cap \bigcup \mathfrak{S}_a \neq \emptyset$, dann
$\exists S \in \mathfrak{S}_a : S \cap X \neq \emptyset$ und $\forall S' \in \mathfrak{S}_a : S' \cap X \neq 0 \Rightarrow S \subseteq S'$.

Die Limes-Annahme besagt soviel wie: Für jede nicht abwegige Bedingung X gibt es eine kleinste Sphäre mit Punkten, die X erfüllen; wir werden gleich darauf zurückkommen. In einer Zeile lassen sich Sphärensysteme also so beschreiben:

$$\{a\} \subseteq S \subseteq S' \subseteq \cdots \subseteq \bigcup \mathfrak{S}_a (\subseteq W).$$

Die Mengeninklusion repräsentiert hier die relative Distanz einer Sphäre vom Zentrum a.[13] Die Menge $\bigcup \mathfrak{S}_a$ aller Punkte, die in irgend einer Sphäre in $\mathfrak{S}_a$ vorkommen, ist die größte Sphäre in $\mathfrak{S}_a$. Was außerhalb von $\bigcup \mathfrak{S}_a$

[13] Man beachte, daß die Bedingungen nicht ausschließen, daß es Sphärensysteme gibt, die die Inklusionskette links noch um $\emptyset \subseteq$ ergänzen. Tatsächlich ist dies in allen Sphärensystemen der Fall, die wir im Abschnitt über die Limes-Annahme (Sla) betrachten werden. Die Hinzunahme der leeren Menge ergibt intuitiv wenig Sinn und ist für die so entstehende Logik folgenlos.

liegt, ist “abwegig”, d.h. eine Möglichkeit, die keinerlei Ähnlichkeit zur Welt im Zentrum aufweist. Wir könnten fordern, daß keine noch so entfernte Möglichkeit abwegig sei. Dann müßten wir Sphärensysteme unter die weitere Bedingung $\bigcup \mathfrak{S}_a = W$ stellen. Zur Grundkonzeption eines Sphärensystems gehört diese Bedingung jedoch nicht.[14]

Die Limes-Annahme (Sla) verhindert, daß die Ketten der Mengeninklusion unendlich dicht werden können – jedenfalls in den Fällen, die uns interessieren. Wenn wir ein Konditional in a beurteilen, dann wollen wir die a-ähnlichste Sphäre finden, in denen das Antezedens möglich ist (d.h. in einigen Welten dieser Sphäre wahr ist). Ohne (Sla) kann aber nun folgendes geschehen: Wir betrachten alle Sphären, in denen A-Welten vorkommen; nennen wir sie A-Sphären. Wenn es zu jeder A-Sphäre eine kleinere (a-ähnlichere) und zu dieser wiederum eine kleinere gibt, ohne daß wir im Fortschreiten von größeren zu kleineren Sphären jemals den Bereich der A-Sphären verlassen müssen, dann gibt es offenbar keine kleinste (ähnlichste) A-Sphäre. Die Bedingung (Sla) schließt genau diese Möglichkeit aus: Für jede nicht abwegige Bedingung X ($X \cap \bigcup \mathfrak{S}_a \neq \emptyset$), gibt es eine X-Sphäre ($S \cap X \neq \emptyset$), welche die kleinste unter den X-Sphären ist ($\forall S' \in \mathfrak{S}_a : S' \cap X \neq 0 \Rightarrow S \subseteq S'$). Die Limes-Annahme ist, bei genauerer Betrachtung, zumindest gewagt. Aber sie vereinfacht die Theorie. Glücklicherweise können wir ohne sie auskommen, wie wir später sehen werden.

Um die Wahrheitsbedingung für Konditionale kurz darstellen zu können, führen wir eine Definition ein: $\min_a(X)$ soll die a naheliegendste Sphäre in $\mathfrak{S}_a$ sein, die sich mit X (nicht-leer) überschneidet, falls es eine solche gibt; anderenfalls sei $\min_a(X)$ die leere Menge. Also

Definition 5. Für $X \cap \bigcup \mathfrak{S}_a = \emptyset$ sei $\min_a(X) = \emptyset$ (in $(W, \mathfrak{S})$). Anderenfalls sei $\min_a(X) = S$ (in $(W, \mathfrak{S})$) gdw

1. $S \in \mathfrak{S}_a$,
2. $S \cap X \neq \emptyset$, und
3. $\forall S' \in \mathfrak{S}_a : S' \cap X \neq \emptyset \Rightarrow S \subseteq S'$.

Schließlich sei für jede Teilmenge X von W, $\mu_a(X) = \min_a(X) \cap X$.

Der Hinweis darauf, daß für jedes $a \in W$ die Abbildungen $\min_a$ und μ_a immer relativ zu einem Sphärenrahmen $(W, \mathfrak{S})$ zu verstehen sind, ist wichtig. Wir können ihn aber im folgenden implizit lassen, da wir uns in einem Kontext immer nur auf einen Sphärenrahmen beziehen werden.

Die erste Zeile der Definition behandelt den Fall, daß X eine abwegige

[14] Ebenso könnte man (Scs) aus der Grundkonzeption von Sphären herausnehmen und durch eine schwächere Zentrierungsbedingung ersetzen. Wir werden darauf zurückkommen; siehe p. 240.

Bedingung ist, d.h. die Menge X nur aus Welten besteht, die außerhalb des Sphärensystems liegen. Im Hauptfall ("anderenfalls ...") einer nicht abwegigen Bedingung X besteht die Menge $\mu_a(X)$ aus allen X-Welten in der kleinsten Sphäre um a, in der es überhaupt X-Welten gibt; mit anderen Worten, $\mu_a(X)$ ist die Menge der a-ähnlichsten X-Welten. Hier setzt die Definition voraus, daß es *genau eine* Sphäre S gibt, welche die Bedingungen 1–3 erfüllt, d.h. eine Sphäre, in denen es X-Welten gibt (Existenz) und die keine kleineren (dem Punkt a näheren) Sphären mit X-Welten enthält (Einzigkeit). In allen *endlichen* Sphärensystemen ist diese Voraussetzung natürlich erfüllt. Für unendliche Sphärensysteme ist es die Limes-Annahme (Sla), welche die Voraussetzung garantiert. Aus der Definition folgt daher unmittelbar, daß

(Sla′) $$X \cap \bigcup \mathfrak{S}_a \neq \emptyset \;\Rightarrow\; \min_a(X) \neq \emptyset.$$

Die Wahrheitsbedingung für $\sqsupset$-Formeln können wir nun so definieren:

($\sqsupset_\ell$) $$a \models A \sqsupset B \text{ gdw } \mu_a[\![A]\!] \subseteq [\![B]\!]$$

(mit $\mu_a[\![A]\!]$ kurz für $\mu_a([\![A]\!])$). Wenn A eine abwegige Bedingung ausdrückt, also $[\![A]\!] \cap \bigcup \mathfrak{S}_a = \emptyset$, dann ist $\mu_a[\![A]\!] = \emptyset$ und somit gilt $a \models A \sqsupset B$ für beliebige Formen B. Im Hauptfall $[\![A]\!] \cap \bigcup \mathfrak{S}_a \neq \emptyset$ läßt sich diese Wahrheitsbedingung z.B. so wie in den Diagrammen 4–7 (s.u.) illustrieren.

In Sphärenmodellen scheitern (Verst), (Trans) und (Kontrap) auf recht anschauliche Weise.

- In den Diagrammen 4–7 liegt $\mu_a(A)$ (mit A^* angedeutet) immer im Bereich der B-Welten; also gilt überall $a \models A \sqsupset B$ nach der Wahrheitsbedingung für ($\sqsupset_\ell$).
- In Diagramm 5 ist jedoch $\mu_a(A \wedge C)$ nicht in B enthalten; also nicht $a \models A \wedge C \sqsupset B$.
- In Diagramm 6 gilt auch $a \models B \sqsupset C$, jedoch nicht $a \models A \sqsupset C$, da $\mu_a(A) \not\subseteq [\![C]\!]$.
- In Diagramm 7 ist $\mu_a(\neg B)$ nicht vollständig in $[\![\neg A]\!] = W \setminus [\![A]\!]$ enthalten; also nicht $a \models \neg B \sqsupset \neg A$.

Von Sphären zu Relationen. In welcher Beziehung stehen Sphärenrahmen zu den früher betrachteten variabel strikten Rahmen für Konditionale? In Sphärenrahmen $(W, \mathfrak{S})$ können wir eine Relation $R^{\mathfrak{S}} \subseteq \wp(W) \times W \times W$ so definieren:

($R^{\mathfrak{S}}$) $$R^{\mathfrak{S}}_X ab \;\text{ gdw }\; b \in \mu_a(X).$$

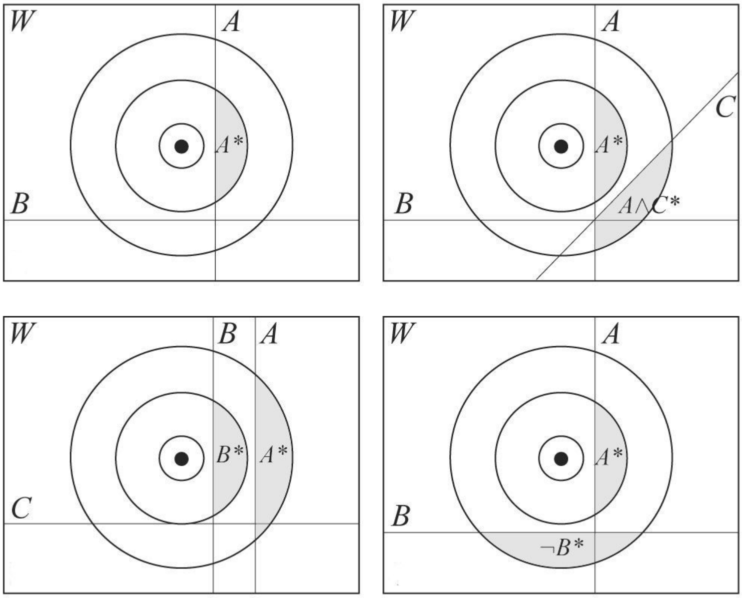


DIAGRAMME 4–7. (Im Uhrzeigersinn.) Der Bereich rechts von der mit A beschrifteten Geraden sei der Bereich der A-Welten. A^* deutet $\mu_a(A)$ an (grauer Bereich); ebenso für die übrigen Buchstaben*. Diagramm 4 illustriert die Wahrheitsbedingung für den Hauptfall $[\![A]\!] \neq \emptyset$. Diagramme 5–7 zeigen jeweils ein Gegenbeispiel zu (Verst), (Trans) und (Kontrap).

(Wir erinnern uns, daß μ_a im Kontext relativ zu $\mathfrak{S}$ ist.) Wenn wir für $R_X ab$ wieder die Notationsvariante $b \in R_X(a)$ verwenden, dann können wir ($R^{\mathfrak{S}}$) auch als Gleichung ausdrücken:

$$R_X^{\mathfrak{S}}(a) = \mu_a(X).$$

Die Relation, so definiert, setzt also zu a alle die Punkte b in Beziehung, welche in der nächstgelegenen Sphäre die Bedingung X erfüllen. Diese Relation ist ersichtlich eine Relation desselben Typs wie die Relation R in variabel relationalen Rahmen (vgl. p. 210).

Aufgrund der Anordnung der Sphären in einem Sphärenrahmen $(W, \mathfrak{S})$ haben die daraus abgeleiteten variabel relationalen Rahmen $(W, R^{\mathfrak{S}})$ gleich eine Reihe von Eigenschaften. Darüberhinaus sind Sphärenrahmen einerseits und daraus abgeleitete relationale Rahmen andererseits im Sinne der zweiten Behauptung des folgenden Satzes äquivalent.

SATZ 5. *Sei* $(W, \mathfrak{S})$ *ein Sphärenrahmen.*

1. *Dann ist* $(W, R^{\mathfrak{S}})$ *ein variabel relationaler Rahmen, der die folgenden Bedingungen erfüllt (statt* $R^{\mathfrak{S}}$ *schreiben wir einfacher* R*):*
   - (id) $R_X ab \Rightarrow b \in X$.
   - (cw) $a \in X \Rightarrow R_X aa$.
   - (mod) $R_X(a) = \emptyset \Rightarrow R_Y(a) \cap X = \emptyset$.
   - (cso) $R_X(a) \subseteq Y$ & $R_Y(a) \subseteq X \Rightarrow R_X(a) = R_Y(a)$.
   - (cs) $a \in X$ & $R_X ab \Rightarrow a = b$.
   - (cv) $R_X(a) \cap Y \neq \emptyset \Rightarrow R_{X \cap Y}(a) \subseteq R_X(a)$.
2. $(W, \mathfrak{S})$ *und* $(W, R^{\mathfrak{S}})$ *sind äquivalent: Für beliebige Formeln* A *und Interpretationen* I,

$$(W, \mathfrak{S}, I) \models A \text{ gdw } (W, R^{\mathfrak{S}}, I) \models A.$$

BEWEIS. *Ad* 1. Wir übersetzen zunächst, nach Def. $(R^{\mathfrak{S}})$, in allen Bedingungen $R_X xy$ als $y \in \mu_x(X)$ $[= \min_x(X) \cap X]$.

(id): Angenommen $X = \emptyset$, dann $\min_a(X) = \emptyset$ und so $\mu_a(X) = \emptyset$. Also gibt es kein b mit $R_X ab$ und somit ist die Implikation (id) auf leere Weise wahr. Im anderen Fall, $X \neq \emptyset$, nehmen wir ferner an, daß $R_X ab$, d.h. $b \in \mu_a(X)$. Dann ist $b \in X$ nach Def. 5.

(cw): Angenommen $a \in X$. Dann ist $X \neq \emptyset$ und also, nach Def. 5 und der Zentrierungsbedingung (Scs) für Sphären, $a \in \min_a(X)$. Also ist $a \in \mu_a(X)$ $[= \min_a(X) \cap X]$, d.h. $R_X aa$.

(mod): Angenommen $\min_a(Y) \cap Y \cap X \neq \emptyset$; zz. $\min_a(X) \cap X \neq \emptyset$. Da $\min_a(Y) \in \mathfrak{S}_a$, ist $\bigcup \mathfrak{S}_a \cap X \neq 0$. Es folgt nach (Sla′), daß $\min(X) \neq 0$ und also (Def. 5.2) auch $\min_a(X) \cap X \neq \emptyset$.

(cso): Wir nehmen an

$$(1)\ \mu_a(X) \subseteq Y \quad \text{und} \quad (2)\ \mu_a(Y) \subseteq X$$

sowie – für *reductio* – $\mu_a(X) \neq \mu_a(Y)$, d.h.

$$\exists b : (3)\ b \in \mu_a(X) \text{ und } (4)\ b \notin \mu_a(Y).$$

Aus (4) folgt $b \notin \min_a(Y)$ oder $b \notin Y$. Letzteres ist jedoch aufgrund von (1,3) ausgeschlossen; also

$$(5)\ b \notin \min_a(Y).$$

Aus (5) und (3) ($\Rightarrow b \in \min_a(X)$) folgt, daß

$$(6)\ \min_a(X) \nsubseteq \min_a(Y).$$

Aus (2) folgt jedoch, daß $S = \min_a(Y)$ eine Sphäre in $\mathfrak{S}_a$ ist mit $S \cap X \neq \emptyset$. Also folgt nach Def 5.3, $\min_a(X) \subseteq S$, im Widerspruch zu (6).

(cs): Wir müssen zeigen, daß im Falle $a \in X$, $\mu_a(X) = \{a\}$. Dazu genügt es nachzuweisen, daß unter der Voraussetzung, $\min_a(X) = \{a\}$. Nach der Zentrierungsbedingung (Scs) ist $\{a\}$ die kleinste Sphäre in $\mathfrak{S}_a$. Da schon diese Sphäre X schneidet, ist $\{a\}$ a-minimal in X.

(cv): Wir nehmen an, daß

$$(1)\ \mu_a(X) \cap Y \neq \emptyset;$$

zz. $\mu_a(X \cap Y) \subseteq \mu_a(X)$, d.h. $\min_a(X \cap Y) \cap X \cap Y \subseteq \min_a(X) \cap X$. Sei $S = \min_a(X)$. Wir zeigen zunächst, daß $S = \min_a(X \cap Y)$. (Da nach (1) X und $X \cap Y$ nicht leer sind, kommt nur der Hauptfall der Def. 5 zur Anwendung.) Aus (1) folgt unmittelbar, daß $S \cap (X \cap Y) \neq \emptyset$. Angenommen nun, es gäbe eine Sphäre $S' \subset S$ mit $S' \cap (X \cap Y) \neq \emptyset$. Dann $S' \cap X \neq \emptyset$. Aber da S minimal in den X-schneidenden Sphären ist, wäre dann, nach Def., 5.3 $S \subseteq S'$ – Widerspruch. Somit haben wir

$$(2)\ \min_a(X) = \min_a(X \cap Y)$$

aus (1) bewiesen. Da nun

$$\min_a(X) \cap X \cap Y \subseteq \min_a(X) \cap X,$$

so folgt durch Einsetzen gemäß (2) die gewünschte Inklusion

$$\min_a(X \cap Y) \cap X \cap Y \subseteq \min_a(X) \cap X.$$

Ad 2. Es genügt zu zeigen, daß

$$\mu_a[\![A]\!] \subseteq [\![B]\!] \text{ in } (W, \mathfrak{S}, v) \text{ gdw } R^{\mathfrak{S}}_{[\![A]\!]} \subseteq [\![B]\!] \text{ in } (W, R^{\mathfrak{S}}, v),$$

was unmittelbar aus $(R^{\mathfrak{S}})$ folgt. ■

Von Relationen zu Sphären. Im vorigen Abschnitt haben wir gezeigt, daß es zu jedem Sphärenrahmen einen daraus abgeleiteten relationalen Rahmen mit den für die Logik **VC** typischen Eigenschaften gibt. Wir zeigen nun, daß auch der umgekehrte Weg möglich ist: Aus relationalen Rahmen können wir Sphärenrahmen ableiten. Das kann natürlich nicht für beliebige relationale Rahmen gelten. So falsifizieren z.B. Rahmen, in denen die Relation nicht die Bedingung (id), $R_X ab \Rightarrow b \in X$, erfüllt, das Schema $A \sqsupset A$, welches in allen Sphärenrahmen gilt. Relationale Rahmen, aus denen sich

Sphärenrahmen ableiten lassen, müssen also eine Reihe von Bedingungen erfüllen. Welches sind diese Bedingungen? Es zeigt sich (im Beweis des nächsten Satzes), daß die folgenden Bedingungen dem Zweck dienen.

(ex) $\exists X : R_X aa$;
(id) $R_X ab \Rightarrow b \in X$;
(cs) $a \in X$ und $R_X ab \Rightarrow a = b$;
(cn) $R_X(a) \cap Y \neq \emptyset \Rightarrow R_Y(a) \subseteq R_X(a)$;
(disj) $R_X(a) \subseteq R_{X\cup Y}(a)$ oder $R_Y(a) \subseteq R_{X\cup Y}(a)$.

Diese Zusammenstellung von Bedingungen hat im jetzigen Zusammenhang vor allem eine technische Bedeutung, indem sie die Wahrheit von Satz 6 garantiert. Wir wollen variabel relationale Rahmen, welche die Bedingungen (ex), (id), (cs), (cn) und (disj) erfüllen, schon im Vorgriff auf Satz 6 *sphäroid* nennen. Für jeden Punkt $a \in W$ eines sphäroiden Rahmens sei eine Menge $\mathfrak{S}_a^R \subseteq \wp(W)$ wie folgt definiert:

($\mathfrak{S}_a^R$) $S \in \mathfrak{S}_a^R$ gdw
1. $\forall b \in W: \; b \in S \Rightarrow \exists X \subseteq W : R_X ab$, und
2. $\forall X \subseteq W: \; X \cap S \neq \emptyset \Rightarrow R_X(a) \subseteq S$.

Die so definierten Mengen sammeln genau die Punkte ein, die zu a ähnlich unter einer (nicht abwegigen) Bedingung X sind. Mit $\mathfrak{S}^R$ bezeichnen wir diejenige Funktion, die jeden Punkt $a \in W$ auf die gerade definierte Menge $\mathfrak{S}_a^R$ abbildet. Es stellt sich heraus, daß die Elemente von $\mathfrak{S}_a^R$ Sphären sind.

Satz 6.
1. *Wenn (W, R) ein sphäroider relationaler Rahmen ist, dann ist $(W, \mathfrak{S}^R)$ ein Sphärenrahmen.*
2. *(W, R) und $(W, \mathfrak{S}^R)$ sind äquivalent, d.h. für beliebige Formeln A und Interpretationen I gilt:*

$$(W, R, I) \models A \text{ gdw } (W, \mathfrak{S}^R, I) \models A.$$

Beweis. *Ad* 1. Wir zeigen, daß $(W, \mathfrak{S}^R)$ die Bedingungen (Scs), (Sli) und (Sla) erfüllt.

(Scs): Mit $\{a\}$ für b in ($\mathfrak{S}_a^R$) eingesetzt, ist die zweite Bedingung der Definition äquivalent zu (cs) und die erste Bedingung folgt unmittelbar aus (ex).

(Sli): Wir nehmen für *reductio* an, daß weder $S \subseteq S'$ noch $S' \subseteq S$ ($S, S' \in \mathfrak{S}_a$). Dann gibt es Punkte b und c mit

$$(1)\ b \in S \text{ und } (2)\ b \notin S' \text{ sowie } (3)\ c \in S' \text{ und } (4)\ c \notin S.$$

Aus (1) folgt nach $(\mathfrak{S}_a^R).1$ $\exists X : R_X ab$, also, nach (id), $b \in X$, und also mit (1), $X \cap S \neq \emptyset$. Daraus schließen wir nach $(\mathfrak{S}_a^R).2$ auf (5) $R_X(a) \subseteq S$. Aus (5) folgt aber auch

$$(*) \qquad R_{(X \cup Y)}(a) \subseteq S.$$

(Denn aus (5) folgt zunächst $R_X(a) \cap S \neq \emptyset$, und da, nach (id), $R_X(a) \subseteq X$, haben wir so $X \cap S \neq \emptyset$. Aber dann auch $(X \cup Y) \cap S \neq \emptyset$, woraus (6) nach $(\mathfrak{S}_a^R).2$ folgt.) Mit einem exakt parallelen Argument führen wir (3) zu der Konklusion

$$(\dagger) \qquad R_{(X \cup Y)}(a) \subseteq S'.$$

Nun wenden wir (disj) an und schließen aus (*) und (†)

$$R_X(a) \subseteq S' \text{ oder } R_Y(a) \subseteq S.$$

Im Falle des linken Disjunkts haben wir dann $b \notin R_X(a)$ aus (2) und im Falle des rechten Disjunkts $c \notin R_Y(a)$ aus (4) – was dem obigen Schluß aus (1) auf $R_X ab$ bzw. dem parallelen Schluß aus (3) auf $R_Y ac$ widerspricht.

(Sla): Wir nehmen an, daß (1) $\bigcup \mathfrak{S}_a^R \cap X \neq \emptyset$. Dann ist nach $(R^{\mathfrak{S}})$ die folgende Familie $\mathcal{X}$ von Mengen S nicht leer:

$$(\mathcal{X}) \qquad \mathcal{X} := \{S \in \mathfrak{S}_a^R : S \cap X \neq \emptyset\}.$$

Wir zeigen, daß $R_X(a) \in \mathcal{X}$ und daß $R_X(a)$ minimal in $\mathcal{X}$ ist.

$R_X(a) \in \mathcal{X}$. Denn Bedingung $(\mathfrak{S}_a^R).1$ ist trivial erfüllt. Die Bedingung $(\mathfrak{S}_a^R).2$ gilt aufgrund von (cn).

$R_X(a)$ ist minimal in $\mathcal{X}$. Denn, sei $S \in \mathcal{X}$. Dann $S \cap X \neq \emptyset$. Es folgt aufgrund $(\mathfrak{S}_a^R).2$, daß $R_X(a) \subseteq S$.

Ad 2. Gegeben eine beliebige Interpretation I, zeigen wir, daß

$$(W, R, I) \models C \text{ gdw } (W, \mathfrak{S}^R, I) \models C$$

durch Induktion über den Aufbau von C. Der interessante Schritt ist der, in dem C ein Konditional $A \sqsupset B$ ist. Hier ist zu zeigen:

$$(*) \qquad \forall b : R_X ab \Rightarrow b \in Y \text{ gdw } \min_a(X) \cap X \subseteq Y.$$

Nun ist $\min_a(X) = S$ genau dann, wenn S minimal in $\mathcal{X}$ (wie soeben in $(\mathcal{X})$ definiert) ist. Wir haben gesehen, daß $S = R_X(a)$. Da, nach (id), $R_X(a) = R_X(a) \cap X$, so ist die rechte Seite von $(*)$ äquivalent zu $R_X(a) \subseteq Y$ und damit auch äquivalent zur linken Seite. ■

Die Limes-Annahme. Wenn die Linie ——————— kürzer als 2 cm wäre, dann wäre sie kürzer, als sie es tatsächlich ist. Wie kurz wäre die Linie, wenn sie kürzer als 2 cm wäre? Wäre sie dann einen halben Zentimeter kürzer? Aber Welten, in denen die Linie ein viertel Zentimeter kürzer wäre, sind unserer Welt ähnlicher als solche, in denen sie einen halben Zentimeter kürzer ist – und um ein achtel Zentimeter kürzere Linien sind unserer Welt noch ähnlicher. Ganz allgemein gibt es für eine Welt, in der die Linie $2-\delta$ cm lang ist, eine unserer Welt ähnlichere, in der sie $2-\frac{\delta}{2}$ cm lang ist. Also gibt es keine Welten, in der die Linie auf die unserer Welt *ähnlichste* Weise kürzer als 2 cm ist. Durch fortgesetzte Verkleinerung der Differenz δ können Welten unserer Welt immer ähnlicher werden und dabei die Bedingung erfüllen, daß die Linie kürzer als 2 cm ist. Es gibt also Antezedens-Bedingungen, die in Weisen erfüllt werden können, die der aktualen Welt immer näher kommen, ohne dabei an eine Grenze zu stoßen. In solchen Fällen ist die Limes-Annahme falsch. Aber wenn die Limes-Annahme für *manche* Bedingungen, die im Antezedens eines Konditionals ausgedrückt werden können, falsch ist, dann kann sie keine Bedingung sein für Modelle, die wir zur Beurteilung *beliebiger* Konditionale verwenden möchten.

Was geschieht, wenn wir auf die Limes-Annahme verzichten? In diesem Fall ist die Funktion $\min_a$ für solche Bedingungen $[\![A]\!]$ nicht definiert, die zu unendlichen Annäherungsketten führen. Daraus folgt, das dann auch der Ausdruck $\mu_a[\![A]\!]$ keine Bedeutung hat. Die rechte Seite der Wahrheitsbedingung

$$(\sqsupset_\ell) \qquad a \models A \sqsupset B \text{ gdw } \mu_a[\![A]\!] \subseteq [\![B]\!]$$

ist dann unwahr – ohne falsch zu sein. Wenn aber die rechte Seite von $(\sqsupset)$ nicht wahr ist, dann kann auch die linke Seite nicht wahr sein, d.h.

$$(\dagger) \qquad a \not\models A \sqsupset B.$$

Also wäre z.B. der Satz "Wenn die Linie [siehe oben] kürzer als 2 cm wäre, dann wäre sie kürzer, als sie es tatsächlich ist" nicht wahr – was offenkundig falsch ist.

Nun ist es so, daß aus der Limes-Annahme folgt, daß jede Teilmenge $\mathcal{A}$ von $\bigcup \mathfrak{S}_a$ eine kleinste und eine größte Sphäre enthält. Da $\mathcal{A}$ verschachtelt unter $\subseteq$ ist (Bedingung Sli), so ist unter dieser Annahme garantiert, daß $\bigcup \mathcal{A}$ und $\bigcap \mathcal{A}$ ebenfalls Sphären in $\bigcup \mathfrak{S}_a$ sind. Ohne die Limes-Annahme müssen wir diese wichtige Garantie in die Definition eines Sphärenrahmens explizit aufnehmen. Ein Sphärenrahmen $(W, \mathfrak{S})$ steht deshalb unter den folgenden Bedingungen:

(Scs) $\mathfrak{S}_a$ ist auf a zentriert: $\{a\} \in \mathfrak{S}_a$.
(Sli) $\mathfrak{S}_a$ ist verschachtelt: $\forall S, S' \in \mathfrak{S}_a : S \subseteq S'$ oder $S' \subseteq S$.
(S∪) $\mathfrak{S}_a$ ist unter Vereinigung abgeschlossen: $\forall \mathcal{A} \subseteq \mathfrak{S}_a : \bigcup \mathcal{A} \in \mathfrak{S}_a$.
(S∩) $\mathfrak{S}_a$ ist unter nicht-leeren Schnitten abgeschlossen:
$\forall \mathcal{A}[\neq \emptyset] \subseteq \mathfrak{S}_a : \bigcap \mathcal{A} \in \mathfrak{S}_a$.

Die Wahrheitsbedingung für ($\sqsupset$) ist nun folgende:

$$(\sqsupset) \quad a \models A \sqsupset B \text{ gdw } \begin{cases} \bigcup \mathfrak{S}_a \cap [A] = \emptyset, \text{ oder} \\ \bigcup \mathfrak{S}_a \cap [\![A]\!] \neq \emptyset \text{ und } \exists S \in \mathfrak{S}_a : S \cap [\![A]\!] \subseteq [\![B]\!]. \end{cases}$$

Die erste Zeile der rechten Seite behandelt den Fall, daß A abwegig ist. Für diesen Fall sieht ($\sqsupset$) vor, daß das Konditional (auf leere Weise) wahr ist; vgl. den Abschnitt über unmögliche Antezendentes, p. 226. Die zweite Zeile behandelt den Hauptfall, $[\![A]\!] \neq \emptyset$. In diesem Fall suchen wir eine Sphäre S, in der es A-Welten gibt. Wenn in allen A-Welten in S auch B wahr ist, dann ist das Konditional wahr und umgekehrt. Das erste Diagramm auf p. 232 illustriert diesen Fall (wobei A^* nun einfach $S \cap [\![A]\!]$ bezeichnet.)

Die Bedingungen ($\sqsupset_\ell$) und ($\sqsupset$) stimmen überein in allen Fällen, in denen die Limes-Annahme erfüllt ist. Für $[\![A]\!] = \emptyset$ ist das offensichtlich. Betrachten wir also den Hauptfall, $[\![A]\!] \neq \emptyset$. Zu zeigen ist

$$\exists S \in \mathfrak{S}_a : S \cap [\![A]\!] \subseteq [\![B]\!] \quad \text{gdw} \quad \min_a [\![A]\!] \cap [\![A]\!] \subseteq [\![B]\!].$$

Unter der Limes-Annahme (Sla) ist $\min_a [\![A]\!] \in \mathfrak{S}_a$ und die Äquivalenz gilt daher von rechts nach links. Für die andere Richtung nehmen wir die linke Seite an. Es sei $\mathcal{A} = \{\mathcal{S} : \mathcal{S} \cap [\![\mathcal{A}]\!] \neq \emptyset\}$. Nach (S∪) ist $\bigcap \mathcal{A} \in \mathfrak{S}_a$ und erfüllt auch die übrigen Bedingungen für $\bigcap \mathcal{A} = \min_a [\![A]\!]$ (Def. 5). Da $\min_a [\![A]\!] \subseteq S$, so auch $\min_a [\![A]\!] \cap [\![A]\!] \subseteq S \cap [\![A]\!]$ und also aufgrund der Annahme über S, $\min_a [\![A]\!] \cap [\![A]\!] \subseteq [\![B]\!]$.

Im Falle unendlicher Ketten von a immer näher kommenden Sphären, ergibt ($\sqsupset$) jetzt das richtige Ergebnis. Das Konditional $A \sqsupset B$ (mit erfüllbarem Antezedenz) ist in a genau dann wahr, wenn folgende Situation zutrifft: Indem wir der Welt a immer näher liegende A-Sphären betrachten, stoßen wir schließlich auf eine A-Sphäre, in der alle A-Welten auch B-Welten sind. Im Falle des Antezedens "Wenn die Linie kürzer als 2cm wäre ..." müssen wir uns von der aktualen Welt a nicht weit entfernen, um auf eine Sphäre von Welten zu stoßen, in der die Linie kürzer ist, als sie es tatsächlich ist. Das Konditional

(47) Wenn die Linie kürzer als 2 cm wäre, dann wäre sie kürzer, als sie es tatsächlich ist

wird also wahr unter der neuen Wahrheitsbedingung ($\sqsupset$). Wie sieht es jedoch mit dem folgenden Konditional aus?

(48) Wenn die Linie kürzer als 2 cm wäre, dann würde der Leser es beim Nachmessen bemerken.

Die Sache ist nicht so klar. Einerseits scheint das Antezedens von (48) von der Art zu sein, welche zur Betrachtung dicht angeordneter Sphären zwingt: Die Differenz δ, um welche die Linie kürzer als 2 cm ist, kann beliebig klein werden. Längen, die man mit einem einfachen Lineal messen kann, sind jedoch nicht beliebig klein. Also ist (48) falsch. Andererseits kann die Rede vom Lineal im Konsequens im passenden Kontext die Limes-Annahme gewissermaßen in Kraft setzen. Ein passender Kontext wäre zum Beispiel der Geometrieunterricht, in dem die Schüler eine Lineal- und Zirkelkonstruktion durch Längenmessung nachprüfen sollen. Dabei werden immer Toleranzen eingeräumt. (Die Differenz δ bekommt einen fixen Minimalwert und kann nur ein ganzzahliges Vielfaches dieses Wertes betragen.) Wenn die Toleranz im Kontext bei 1 mm liegt, dann gelten alle Linien als gleich lang, wenn sie sich nicht um mehr als 1 mm unterscheiden. Das hat den Effekt, daß "kürzer als" in (48) nicht als eine dichte, sondern als eine diskrete Relation interpretiert wird. Unter dieser Interpretation sind für die Beurteilung von (48) nur endlich viele Sphären zu betrachten und das Konditional ist wahr. Ob wir ein Konditional in dichten oder diskreten Sphärensystemen beurteilen sollten, hängt offenbar nicht nur vom semantischen Gehalt des Antezedens, sondern auch vom Kontext ab.

Das ändert nichts daran, daß die Limes-Annahme etwas Falsches über die Relation der Ähnlichkeit zwischen Welten sagt. Da Ähnlichkeit grundsätzlich eine dichte Relation ist, so sind Sphärensysteme immer dicht angeordnet. Aber in vielen Fällen erlaubt der Kontext die Dichte zu ignorieren. In solchen Fällen sind die relevanten Sphären diskret voneinander unterschieden. Für die relevanten Sphären gilt dann die Limes-Annahme, obwohl das System als solches dicht angeordnet ist.

Logik der Sphären. Die Limes-Annahme spielt für die logische Theorie keine Rolle; sie ist für dichte und diskrete Sphärensysteme gleich. Es sei SPH_ℓ die Klasse aller Sphärenmodelle so, wie wir diese zu Anfang, d.h. mit der Limes-Annahme (Sla) und der Wahrheitsbedingung ($\sqsupset_\ell$) definiert haben. Dagegen sei SPH die Klasse der Sphärenmodelle so, wie wir sie im letzten Abschnitt neu definiert haben, d.h. ohne (Sla), dafür jedoch mit dem Abschluß unter Vereinigung und Schnitt und der neuen Wahrheitsbedingung ($\sqsupset$), welche die Limes-Annahme nicht voraussetzt. Dann sind in SPH_ℓ und in SPH genau dieselben Formeln gültig. Es stellt sich heraus, daß dies die

Das System **VC**

τ.	Alle Tautologien
REA.	$\dfrac{A \leftrightarrow A'}{(A \sqsupset B) \to (A' \sqsupset B)}$
RK.	$\dfrac{B_1 \wedge \cdots \wedge B_n \to B}{(A \sqsupset B_1) \wedge \cdots \wedge (A \sqsupset B_n) \to (A \sqsupset B)}$ $(0 \leq n)$
MP.	$\dfrac{A, A \to B}{B}$
ID.	$A \sqsupset A$
CSO.	$(A \sqsupset A') \wedge (A' \sqsupset A) \to ((A \sqsupset B) \to (A' \sqsupset B))$
CV.	$(A > B) \wedge (A \sqsupset C) \to (A \wedge B \sqsupset C)$
MOD.	$(\neg A \sqsupset \bot) \to (B \sqsupset A)$
CW.	$(A \sqsupset B) \to (A \to B)$
CS.	$A \wedge B \to (A \sqsupset B)$

Theoreme der Logik **VC** sind; vgl. die Tafel unten.[15]

Bis einschließlich MOD sind alle Schemata der Tafel in die Sphärensysteme gewissermaßen eingebaut. Wenn die semantische Analyse kontrafaktischer Konditionale in Sphärensystemen überzeugt, dann kann keines dieser Schemata strittig sein. Etwas anders steht es um das Schema CS. Es wird durch die starke Zentrierungsbedingung (Scs) erzwungen, wonach in jedem Sphärensystem auf einer Welt a die Menge $\{a\}$ als kleinste Sphäre enthalten ist. Einerseits ist die Bedingung durch die Modellierungsidee sehr gut motiviert; andererseits ist das zugehörige Schema negativ auffällig: Nicht jedes zufällig wahre Paar von Aussagen scheint entsprechende Konditionale (in beide Richtungen) zu implizieren. Vielleicht sollten wir deshalb einer

[15] Ein Vollständigkeitsbeweis über die Methode der kanonischen Modelle findet sich in [182].

schwächeren Zentrierungsbedingung den Vorzug geben: $\{a\}$ ist nicht immer die kleinste Sphäre in $\mathfrak{S}_a$, aber a ist der gemeinsame Kern aller Sphären auf a, d.h. a ist in allen Sphären als Element enthalten. Genauer formuliert, ist diese Bedingung der schwachen Zentrierung gemeint:[16]

(Scw) $\exists S \in \mathfrak{S}_a : S \neq \emptyset$ und $\forall S \in \mathfrak{S}_a : S \neq \emptyset \Rightarrow a \in S.$

Diese Bedingung entspricht dem unkontroversen Schema CW. Wenn wir aus den Schemata der Tafel das Schema CS streichen, dann erhalten wir eine Axiomatisierung der in schwach zentrierten Sphärensystemen gültigen Formeln. Dieses System wird **VW** genannt. Für eine Axiomatisierung von **VC** ist das Schema CW in der Tafel redundant; vor dem Hintergrund der anderen Schemata folgt dieses aus CS. Fügen wir **VC** noch das Schema

CEM. $(A \sqsupset B) \vee (A \sqsupset \neg B)$

hinzu, so entsteht Stalnakers System **C2** (bei Lewis **VCS**). In der Sphärensemantik entspricht CEM der Bedingung, daß es in jedem Sphärensystem für jede nicht abwegige Proposition X eine Sphäre S gibt, so daß $X \cap S$ genau eine Welt enthält.

* * *

Die Theorie kontrafaktischer Konditionale, die wir bis hierher vorgestellt haben, ist eine reife Theorie, nicht anders als die Modelltheorie der Modallogik. Die Theoriebildung hat schon früh einen Fixpunkt erreicht – im Grunde schon mit Lewis' Untersuchung [182] aus dem Jahre 1973.[17] Das gilt sowohl für die Art der semantischen und logischen Analyse als auch für die ausgezeichnete Stellung gewisser Systeme unter der Grundinterpretation. Die Systeme **VW** und **VC** sind – wie **S4** und **S5** in der Modallogik – als Referenzpunkte wenig umstritten. Weitgehend unumstritten ist es auch, daß kontrafaktische Konditionale wichtige Aspekte der Welt beschreiben. Daß Kochsalz löslich ist, ist eine Tatsache, und genau diese Tatsache wird auch beschrieben durch das Konditional "Wenn man Kochsalz in Wasser geben würde, dann würde es sich darin auflösen". Mit anderen Worten,

[16] Wenn wir (Scs) durch die neue Bedingung ersetzen, daß a in allen Sphären enthalten sein muß, dann garantiert jetzt nur die Bedingung (S$\cup$), daß ein Sphärensystem nicht leer ist. Aber die bloße Garantie, daß $\emptyset \in \mathfrak{S}_a$, ist zu wenig für eine nicht-triviale logische Theorie. Deshalb brauchen wir eine Bedingung, die mindestens eine nichtleere Sphäre in $\mathfrak{S}_a$ erzwingt. Genau das ist der Sinn des ersten Konjunkt von (Scw).

[17] Diese schließt an noch frühere Arbeiten an von Lennart Åquvist, Kit Fine, Hans Kamp und Robert Stalnaker.

wenn Dispositionseigenschaften auf Stoffe zutreffen können – was kaum jemand bestreitet –, dann müssen auch die entsprechenden kontrafaktischen Konditionale wahr sein können.

Die Theorie indikativischer Konditionalsätze, der wir uns jetzt zuwenden, ist dagegen weitaus unabgeschlossener. Das betrifft auch die grundsätzliche Frage, ob sich für solche Konditionale überhaupt Bedingungen angeben lassen, unter denen sie etwas zutreffend beschreiben, also wahr sind. Unumstritten ist, daß ein Sprecher mit einem Konditional wie "Wenn der Zug pünktlich abgefahren ist, dann wird er auch pünktlich eintreffen" Auskunft über seinen Überzeugungszustand gibt, nämlich, daß dieser so ist, daß der Sprecher glaubt, der Zug komme pünktlich an, sobald er glaubt, daß er pünktlich abgefahren sei. Wenn sein Überzeugungszustand nicht so ist, dann sollte er das Konditional nicht behaupten. Daß zu Konditionalen Bedingungen gehören, unter denen sie richtig behauptet werden können, ist also nicht kontrovers. Aber können Konditionale darüber hinaus auch wahr oder falsch sein? Beschreiben sie Tatsachen, so wie es das kontrafaktische Kochsalz-Konditional tut? Wenn ja, welche Tatsache wird durch diese eigentümliche Kombination zweier Sätze beschrieben?

6. Indikativische Konditionale

Jacksons Tatsächlichkeitsargument (vgl. p. 201) weist auf einen wesentlichen Unterschied zwischen kontrafaktischen und indikativischen Konditionalen hin: Jene bringen alternative mögliche Welten ins Spiel, diese tun das nicht. Wenn indikativische Konditionale überhaupt Wahrheitsbedingungen haben, dann spielen für die Wahrheit von *Wenn A, dann B* in einer Welt a allein Eigenschaften dieser Welt a eine Rolle. Wenn die Wahrheitsbedingung in irgendeiner Weise rekursiv sein soll, dann bedeutet das, daß der Wahrheitswert von *Wenn A, dann B* in a bestimmt wird von Eigenschaften der Teilsätze A und B in a. Die einfachste und daher zunächst auch plausibelste Hypothese ist, daß die einzig relevanten Eigenschaften, um die es hier gehen kann, die Wahrheitswerte von A und B in a sind. Nach dieser Hypothese ist das Konditional *Wenn A, dann B* eine Wahrheitsfunktion der Teilsätze A und B, und offenbar kommt hier nur diejenige Funktion in Frage, die wir so ausdrücken können:

MAT. $\quad a \models \textit{Wenn } A, \textit{ dann } B \;$ gdw $\; a \not\models A$ oder $a \models B$,

d.h. indikativische Konditionale haben die Wahrheitsbedingungen materialer Konditionale.

Diese Hypothese MAT steht, für sich betrachtet, nicht gut da angesichts von Beispielen wie diesen:

(49) (a) Peter hat einige Klausuren bestanden.
(b) Peter wird zur Abschlußprüfung zugelassen.
$^{?}$*Also*: (c) Wenn Peter keine Klausur bestanden hat, dann wird er zur Abschlußprüfung zugelassen.

(50) (a) Die Staatsschulden werden nicht abgebaut.
(b) Die Krise hält an.
$^{?}$*Also*: (c) Wenn die Schulden abgebaut werden, dann hält die Krise an.

Unter den Annahmen (a) und (b) würden wir die konditionale Konklusion (c) normalerweise nicht behaupten wollen. Aber (c) ist wahr unter den Annahmen (a) und (b) und der Wahrheitsbedingung MAT. Wir müssen also entweder die Hypothese MAT aufgeben oder den Schluß von der Wahrheit auf die Behauptbarkeit von (c) blockieren.

Grice: Konditionale und Gesprächsimplikatur. Nach Grice [118] sollten wir bei der Betrachtung von Konditionalen pragmatische Aspekte mit einbeziehen. Es ist eine Binsenweisheit, daß man nicht allein durch das Äußern wahrer Sätze zu einem gelungenen Gespräch beiträgt. Beispielsweise ist es nicht sehr hilfreich, wenn ein Arzt seinem Patienten mitteilt, daß er unter Migräne oder an einem Hirntumor leidet, wenn der Arzt entweder die Migräne oder den Tumor schon festgestellt hat. Die schwächere disjunktive Aussage ist zwar wahr, aber im Kontext interessiert allein die informativere, stärkere Aussage. Wahre Aussagen sind nicht in jedem Kontext äußerbar oder *behauptbar*, wie wir fortan sagen wollen. [18] Für ein erfolgreiches Gespräch gibt es gewisse Regeln. Gespräche sind kooperative Unternehmungen, die vornehmlich dem Zweck dienen, Information auszutauschen. Daher gilt die oberste Regel, sich im Hinblick auf den Zweck kooperativ zu verhalten. Das ist man nur dann, wenn man ehrlich ist, sich so deutlich wie möglich ausdrückt, zum Punkt spricht und nicht hinter dem Berg hält. Grice hat diese Bedingungen in Maximen gefaßt, zu denen auch die folgende gehört: [19]

---

[18] Umgekehrt, müssen auch nicht alle behauptbaren Aussagen wahr sein. Wenn der Sprecher seinen “epistemischen Pflichten” nachkommt – Aussagen nach bestem Wissen und Gewissen macht –, dann können wir ihn nicht für die Behauptung von Aussagen kritisieren, die sich als falsch herausstellen.

[19] Die anderen Maximen sind die Maxime der *Qualität* (behaupte nur, was Du für wahr und ausreichend begründet hältst), die Maxime der *Relation* (enthalte Dich irrelevanter Behauptungen), und die Maxime der *Art und Weise* (drücke Dich so klar und deutlich wie möglich und in guter Reihenfolge aus). Daß Gespräche im Normalfall kooperative Unternehmungen sind und also von Normen bestimmt werden, die den Erfolg der Kooperation sichern sollen, ist einleuchtend. Die Theorie von Grice in [118] ist ein erster Versuch, diese Normen zu bestimmen und die Unterscheidung zwischen Wahrheit und Behaupt-

Die Maxime der Quantität

Behaupte nicht weniger als Du weißt und nicht mehr als dem Zweck des Gesprächs dienlich ist.

Normalerweise unterstellen Sprecher einander, daß sie die Maxime befolgen. In diesem Sinne trägt jeder Redebeitrag eine sogenannte *Gesprächsimplikatur*: In einem Gesprächskontext ist eine Aussage B Implikatur einer Behauptung A (unter der Quantitätsmaxime), wenn

(a) der Hörer davon ausgehen darf, daß der Sprecher, indem er A behauptet, die Maxime der Quantität befolgt;
(b) diese Annahme den Hörer berechtigt auf B zu schließen (und dieser dazu auch in der Lage ist); und
(c) der Sprecher davon ausgehen muß, daß (b) der Fall ist.

Nach der Maxime der Quantität trägt Peters Äußerung des Satzes "Die Eintracht oder die Borussia spielen um den Pokal" in einem typischen Kontext die Implikatur, daß Peter sich nicht in der Lage sieht, zu sagen, welche von beiden Mannschaften ins Finale einziehen wird. Diese Implikatur kann er aufheben, indem er dem Satz einfach hinzufügt: "... und ich weiß auch schon, wer es sein wird". Die Möglichkeit, eine Implikatur B einer Äußerung von A zu streichen, indem man einfach $\neg B$ der Aussage A hinzufügt, unterscheidet Implikaturen von Implikationen – und ist somit ein Test auf das Vorliegen einer Implikatur. Denn, wenn A die Aussage B (material) impliziert, dann ist die Konjunktion $A \wedge \neg B$ widersprüchlich. Wenn die Behauptung von A dagegen B "implikiert", dann generiert die Hinzufügung von $\neg B$ keinen Widerspruch, sondern "löscht" einfach nur die normalerweise unterstellte Aussage B.

Die Maxime der Quantität gibt eine Möglichkeit an die Hand, zu erklären, was Beispiele wie (49) und (50) so merkwürdig macht, ohne daß wir die einfache Wahrheitsbedingung MAT für indikativische Konditionale aufgeben müssen. Es läßt sich keine einigermaßen normale Situation denken, in der ein Sprecher (c) zurecht behaupten könnte, wenn er (a) oder (b) für wahr hält. Denn, in normalen Kontexten geäußert, trägt (c) die Implikatur, daß der Sprecher nichts stärkeres behaupten kann. Aber genau diese Implikatur ist falsch; er könnte ja (a) oder (b) behaupten und damit dem Gesprächszweck besser dienen. Die Bedingungen für die Behauptbarkeit von (c) sind also nicht erfüllt. Das Konditional (c) ist, obgleich wahr unter einer der Voraussetzungen (a) oder (b), nicht behauptbar. Daher würde

barkeit systematisch anzugehen. Bei genauerem Hinsehen steht eine solche Theorie vor schwierigen Herausforderungen; viele davon behandelt schon Grice in späteren Schriften (gesammelt in [119]); vgl. auch [60].

auch niemand in einem normalen Gespräch ein Argument wie (49) oder (50) vorlegen wollen, obgleich es doch gültig ist.

Die Theorie der Gesprächsmaximen blockiert also den Schluß von der Wahrheit eines Konditionals auf dessen Behauptbarkeit. Sie erklärt, warum wir an den Konditionalen (c) in (49) und (50) Anstoß nehmen: Die Konditionale sind wahr, jedoch falsch verwendet. Die andere Möglichkeit, unserer Ablehnung der Konditionale zu erklären, bestünde in einem Nachweis ihrer Unwahrheit: Wir würden so die Wahrheitsbedingung MAT zurückweisen und (c) als Beispiele einer neuen Sorte von Konditionalen identifizieren. Welcher Erklärung sollten wir den Vorzug geben?

Gesprächsmaximen haben eine einfache Erlärung im sozialen Zweck von Sprache überhaupt. Menschen sind auf Kooperation angewiesen. Gesprächsmaximen sind Normen, die den kooperativen Zweck sprachlicher Kommunikation sichern sollen. Grice zieht daraus einen methodologischen Schluß, der manchmal das "Grice'sche Rasiermesser" genannt wird:[20] Wenn zwei Erklärungen eines sprachlichen Phänomens angeboten werden, von denen die eine ausschließlich auf allgemeine Kooperationsmaximen zurückgreift, während die andere spezifischere Erklärungen bemühen muß, dann ist immer der ersten der Vorzug zu geben, da sie auf einer allgemeineren und letztlich sprachunabhängigen Erklärungsbasis beruht.

* * *

Die entscheidende Frage ist: Reichen Gesprächsmaximen aus, um die gegen die Hypothese MAT sprechenden Daten über indikativische Konditionale zu erklären? Der Versuch, MAT durch den Rückgriff auf Gesprächsmaximen zu retten, ist von zwei Möglichkeiten bedroht. Erstens, könnte die Theorie material wahre Konditionale als behauptbar passieren lassen, die es tatsächlich nicht sind. Die Grice'sche Theorie würde dann nicht alle Gegenbeispiele zu MAT ausschalten können – ihr Behauptbarkeitsbegriff würde "übergenerieren". Zweitens, könnte die Theorie auch untergenerieren. Dann würde sie material wahre Konditionale als nicht behauptbar aussondern, obwohl sie durchaus behauptbar sind. Die folgenden Beispiele bedrohen die Theorie mit dem Vorwurf der Untergenerierung:

(51) Es wird nicht regnen. Falls es (doch) regnet, gehen wir ins Haus.
(52) Das Spiel gewinnen wir. (Selbst) wenn Meier ausfällt, gewinnen wir.

[20] Das spielt auf Ockhams Rasiermesser in der Ontologie an: Entitäten sind nicht ohne Not zu vermehren. Das Grice'sche Rasiermesser wird gegen die Vermehrung semantischer Entitäten geführt.

Die Beispiele sind analog zu (49) und (50) konstruiert. Der jeweilige Nachsatz ist ein Konditional; in (51) wird dessen Antezedens eingangs verneint, bzw. in (52) wird dessen Konsequens eingangs bejaht. Im Gegensatz zu (49) und (50) haben diese Paare aber überhaupt nichts merkwürdiges an sich. Wir können uns zwanglos Situationen vorstellen, in denen jeweils einer der Sätze in (51) bzw. (52) behauptet werden. In solchen Situationen behauptet der Sprecher ein Konditional, obwohl er zuvor eine stärkere Behauptung macht. Damit verletzt der Sprecher die Quantitätsimplikatur des Konditionals – und wir würden das nicht beanstanden. [21]

Angenommen Peter ist überzeugt, daß die Eintracht gewinnt, daß der Torwart nicht ausfällt und der Einsatz von Meier, im Gegensatz zu dem des Torwarts, keinen Einfluß auf das Spielergebnis haben wird. Nach der Grice'schen Theorie sind beide Konditionale nicht behauptbar:

(53) Wenn Meier ausfällt, gewinnt die Eintracht.
(54) Wenn der Torwart ausfällt, gewinnt die Eintracht.

Aber während Peter sehr wohl und zurecht (53) behaupten könnte, würde er (54) sicher nicht behaupten wollen. Den Unterschied in der Behauptbarkeit der beiden Konditionale können wir nicht allein im Rückgriff auf die Quantitätsmaxime erklären.

Daß die Quantitätsmaxime allenfalls *cum grano salis* zu nehmen ist, zeigt sich selbst an ihrem Schaufensterstück, dem Gebrauch von Disjunktionen. Angenommen Peter ist sich sicher, daß Meier spielen und daß Müller aufgrund einer schweren Verletzung ausfallen wird. Nach der Maxime sollte Peter deshalb auf die Frage nach der Mannschaftsaufstellung nicht mit der Disjunktion "Meier oder Müller spielt" antworten. Angenommen aber, Peter und Hans sind im Stadion, die Aufstellung wird kurz auf dem Bildschirm gezeigt, Hans ist abgelenkt, und Peter sieht nur den Anfangsbuchstaben "M". Hans fragt: "Wer spielt?" Die Antwort "Meier oder Müller" scheint völlig in Ordnung zu sein, obschon sich an der doxastischen Situation von Peter nichts geändert hat. [22] Peter könnte – müßte aber nicht – so fortset-

---

[21] Gegen (51) könnte eingewandt werden, daß das Konditional mehr als eine bloße Abschwächung von "Es wird nicht regnen" ausdrückt. Es scheint auch eine *Aufforderung* auszudrücken, welche die bloße Negation des Antezedens nicht zu verstehen gibt. In (52) scheint das Konditional nicht eine Abschwächung des Konsequens, sondern, im Gegenteil, eine Verstärkung auszudrücken: Wir gewinnen *in jedem Fall*. Es ist nicht zu bestreiten, daß die Grice'sche Theorie über beträchtliche Ressourcen verfügt. Die Herausforderung besteht darin, diese Ressourcen in einer Weise zu mobilisieren, die nicht *ad hoc* erscheinen muß und dem Grice'schen Rasiermesser Scharten zufügt. Wie groß – und vielleicht unüberwindlich – die Schwierigkeiten wirklich sind, zeigt [146].

[22] Beispiel adaptiert aus [148, p. 23]. Das Beispiel weist voraus auf den Begriff der Robustheit einer Information; siehe unten. Die schwächere Auskunft "Meier oder Müller" ist robust gegen die Möglichkeit, daß Meier nicht spielt, auch wenn Peter das für wenig

zen: "... Aber Müller ist verletzt. Also spielt Meier." Ein solcher, völlig natürlicher Gesprächsbeitrag, würde durch die Quantitätsmaxime infrage gestellt. Wir brauchen daher eine alternative Theorie des angemessenen Gebrauchs solcher Sätze.

7. Konditionale und konditionale Wahrscheinlichkeit

Eine Beobachtung, die uns auf den richtigen Weg bringen könnte, ist, daß Peter (53) eine höhere *Wahrscheinlichkeit* zumißt als (54). Unter den gemachten Annahmen, ist Peters Wahrscheinlichkeit für das erste Konditional sehr hoch, während die für das zweite Konditional sehr niedrig ist. Peter sollte sicher nichts behaupten, was er für wenig wahrscheinlich hält. Und, soweit relevant und schicklich, sollte er Aussagen zum Gespräch beisteuern, von deren hoher Wahrscheinlichkeit er überzeugt ist. Es ist daher eine naheliegende Frage, ob und, wenn ja, wie die Behauptbarkeit eines Konditionals mit der Wahrscheinlichkeitseinschätzung des Sprechers einhergeht.

Exkurs über Wahrscheinlichkeit. Wahrscheinlichkeiten werden Ereignissen zugeordnet – das wollen wir hier jedenfalls annehmen. Wir wollen ferner annehmen, daß Ereignisse durch die Sätze einer Sprache beschrieben werden. Eine Wahrscheinlichkeitsfunktion über einer Menge von Ereignissen ist dann eine Abbildung p, die jedem Satz A, der ein Ereignis in der Menge beschreibt, einen Wert $\mathrm{p}(A)$ zwischen einem Minimum und einem Maximum zuordnet; konventionellerweise legt man das reelle Werteintervall $[0,1] = \{n \in \mathbb{R} : 0 \leq n \leq 1\}$ zugrunde.[23] Von einer *Wahrscheinlichkeitsfunktion* $\mathrm{p} : \mathrm{FML} \longrightarrow [0,1]$ erwarten wir, daß sie eine Reihe von Bedingungen erfüllt. So soll es nicht darauf ankommen, wie wir Ereignisse beschreiben; logisch äquivalente Beschreibungen sollen gleiche Wahrscheinlichkeitswerte erhalten:

P0. $$\mathrm{p}(A) = \mathrm{p}(B), \text{ falls } A \equiv B.$$

Tautologien sind absolut sichere Ereignisse und erhalten deshalb den Maximalwert 1 (Bedingung der *Normalität*):

P1. $$\mathrm{p}(\top) = 1.$$

wahrscheinlich hält. Manchmal kann in einem Gespräch Robustheit wichtiger sein als die riskantere Erfüllung der Quantitätsmaxime.

[23] Wenn die Anzahl der Atome der zugrundeliegenden Sprache endlich ist, d.h. wenn es in der Situation, die wir modellieren möchten, nur endlich viele Ereignissen zu betrachten gibt, dann tut es das Interval $[0,1]$ in den rationalen Zahlen $\mathbf{Q}$ genauso gut.

Wenn zwei Ereignisse einander ausschließen, d.h. die Wahrscheinlichkeit ihres gemeinsamen Auftretens gleich Null ist, dann ist die Wahrscheinlichkeit, daß das eine oder das andere stattfindet, gleich der Summe der Einzelwahrscheinlichkeiten. Dies ist die Bedingung der *endlichen Summierbarkeit*:

P2. $$\mathrm{p}(A \vee B) = \mathrm{p}(A) + \mathrm{p}(B), \text{ falls } \mathrm{p}(A \wedge B) = 0.$$

Damit sind die wesentlichen Eigenschaften von Wahrscheinlichkeitsfunktionen $\mathrm{p} : \mathrm{FML} \longrightarrow [0,1]$ beschrieben.[24]

Da A und $\neg A$ einander ausschließen, dürfen wir nach P2 auf $\mathrm{p}(A \vee \neg A) = \mathrm{p}(A) + \mathrm{p}(\neg A)$ schließen. Da $A \vee \neg A$ eine Tautologie ist, so wissen wir nach P1 (und P0), daß $\mathrm{p}(A \vee \neg A) = 1$ und also $\mathrm{p}(A) + \mathrm{p}(\neg A) = 1$, woraus unmittelbar folgt, daß

P3. $$\mathrm{p}(\neg A) = 1 - \mathrm{p}(A).$$

Wenn wir in diese Gleichung $\top$ für A einsetzen, dann folgt $\mathrm{p}(\neg\top) = 0$, d.h. Widersprüche $\perp$ $(= \neg\top)$ haben den minimalen Wahrscheinlichkeitswert 0.

Die Addition von Wahrscheinlichkeiten können wir auch auf "allgemeinere" Weise beschreiben, d.h. ohne die Bedingung, daß A und B einander ausschließen:

P2′. $$\mathrm{p}(A \vee B) = \mathrm{p}(A) + \mathrm{p}(B) - \mathrm{p}(A \wedge B).$$

Tatsächlich sind P2 und P2′ äquivalent. Der Schluß von P2′ auf P2 ist trivial. Für die andere Richtung stellen wir zunächst fest, daß

$$\text{(a) } A \vee B \equiv A \vee (\neg A \wedge B) \quad \text{und} \quad \text{(b) } B \equiv (A \wedge B) \vee (\neg A \wedge B).$$

Also gelten nach P0 die entsprechenden Wahrscheinlichkeitsgleichungen. Da A und $\neg A \wedge B$ einander ausschließen, folgt nach P2 aus (a), daß

(a′) $$\mathrm{p}(A \vee B) = \mathrm{p}(A) + \mathrm{p}(\neg A \wedge B).$$

---

[24] Statt Wahrscheinlichkeitsfunktionen unter diese Postulate zu stellen, können wir die Postulate auch aus einer etwas anderen Definition herleiten. Dazu gehen wir aus von einer beliebigen nichtleeren und endlichen Menge Ω möglicher *Ergebnisse*. Über Ω verteilen wir mit einer Funktion $d : \Omega \longrightarrow [0,1]$ Werte so, daß diese zu 1 summieren, d.h. $\Sigma\{d(x) : x \in \Omega\} = 1$. Nun betrachten wir den *Ereignisraum* $\wp(\Omega)$ und definieren die Wahrscheinlichkeit $\mathrm{p}(E)$ für jedes Ereignis $E \subseteq \Omega$ durch $\mathrm{p}(E) = \Sigma\{d(x) : x \in E\}$. Dann gilt $\mathrm{p}(E_1 \cup E_2) = \mathrm{p}(E_1) + \mathrm{p}(E_2)$, falls $E_1 \cap E_2 = \emptyset$. Ferner gilt $\mathrm{p}(E \cup (\Omega \setminus E)) = \mathrm{p}(\Omega) = 1$.

Auch auf das einander ausschließende Paar $A \wedge B$ und $\neg A \wedge B$ können wir P2 anwenden und erhalten so aus (b),

$$\text{(b}'\text{)} \qquad \mathrm{p}(B) = \mathrm{p}(A \wedge B) + \mathrm{p}(\neg A \wedge B),$$

was wir zu

$$\mathrm{p}(\neg A \wedge B) = \mathrm{p}(B) - \mathrm{p}(A \wedge B)$$

umstellen können. Durch Einsetzen in (a′) folgt nun die gewünschte Gleichung P2′.

Man kann sich schnell davon überzeugen, daß Wahrheitsbewertungen $[\![\ ]\!]_I : \mathrm{FML} \longrightarrow \{0, 1\}$ Grenzfälle von Wahrscheinlichkeitsbewertungen sind.[25] Das ist schon intuitiv zu erwarten und wird durch die bisher betrachteten P-Bedingungen bestätigt. Man beachte aber auch, daß die Wahrscheinlichkeit einer wahrheitsfunktional zusammengesetzten Aussage im allgemeinen keine Funktion der Wahrscheinlichkeiten ihrer Teilaussagen ist. Wahrheitsfunktionen müssen keine Wahrscheinlichkeitsfunktionen sein. Die Negation ist, nach P3, wahrscheinlichkeitsfunktional. Die Disjunktion ist es jedoch nicht. Denn $\mathrm{p}(A) = \mathrm{p}(A')$ garantiert nicht, daß $\mathrm{p}(A \vee B) = \mathrm{p}(A' \vee B)$. Würfel haben eine Chance von 1/2 eine gerade (A) oder ungerade (A') Zahl zu zeigen. Ferner ist die Chance, eine gerade Zahl oder eine Sechs zu zeigen ($A \vee B$) gleich der Wahrscheinlichkeit eine gerade Zahl zu zeigen – welche (um 1/6) geringer ist als die Wahrscheinlichkeit, eine ungerade Zahl oder eine Sechs zu zeigen ($A' \vee B$).

Es gilt ferner aufgrund der Bedingungen:

$$\text{PL.} \qquad \text{Wenn } A \models B, \text{ dann } \mathrm{p}(A) \leq \mathrm{p}(B).$$

Denn $\mathrm{p}(A \vee \neg B) \leq 1$, während $\mathrm{p}(A \vee \neg A) = 1$; also

$$(*) \qquad \mathrm{p}(A \vee \neg B) \leq \mathrm{p}(A \vee \neg A).$$

Nun folgt aus $A \models B$, daß $\mathrm{p}(A \wedge \neg B) = 0$. Also dürfen wir P2 anwenden und erhalten aus (∗),

$$\mathrm{p}(A) + \mathrm{p}(\neg B) \leq \mathrm{p}(A) + \mathrm{p}(\neg A).$$

Aber dann haben wir $\mathrm{p}(\neg B) \leq \mathrm{p}(\neg A)$ und also nach P3, $\mathrm{p}(A) \leq \mathrm{p}(B)$.

[25] Unter einer Wahrheitsbewertung $[\![\ ]\!]_I$ wollen wir, wie schon zuvor, die Erweiterung einer Interpretation I der Atome in $\{0, 1\}$ verstehen, welche in Einklang mit den bekannten Wahrheitsbedingungen steht; d.h. $[\![\neg A]\!]_I = 1 - [\![A]\!]_I$, $[\![A \wedge B]\!]_I = 1$ gdw $[\![A]\!]_I = 1 = [\![B]\!]_I$ usw.

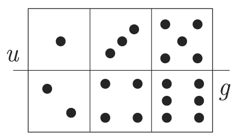
Oft interessiert uns nicht die Wahrscheinlichkeit schlechthin eines Ereignisses, sondern die Wahrscheinlichkeit des Ereignisses unter der Voraussetzung, daß andere Ereignisse eintreffen, d.h. unter einer Bedingung B. In diesem Fall fragen wir nicht mehr nach den Wahrscheinlichkeiten von Ereignissen im ursprünglichen, "großen" Raum von Möglichkeiten, sondern nach Wahrscheinlichkeiten in einem kleineren, auf B eingeengten Raum. Anders gesagt: Wir gehen über von einer Wahrscheinlichkeitsfunktion zu einer anderen. Dieser Übergang ist natürlich nicht willkürlich, sondern wird bestimmt durch die bereits vergebenen Wahrscheinlichkeiten im ursprünglichen Ereignisraum. Das drücken wir durch eine definierende Gleichung aus:[26]

Def. p_{B} $$\mathrm{p}_B(A) = \frac{\mathrm{p}(A \wedge B)}{\mathrm{p}(B)}, \text{ wenn } \mathrm{p}(B) \neq 0.$$

Die Funktion p_B mißt die durch B *bedingte Wahrscheinlichkeit* eines Ereignisses. Man beachte, daß die Funktion $\mathrm{p}_B(A)$ nur definiert ist für den Fall, daß B eine Möglichkeit beschreibt.[27]

Aus der Definition folgt unmittelbar eine Wahrscheinlichkeitsbedingung für Konjunktionen, nämlich

P4. $$\mathrm{p}(A \wedge B) = \mathrm{p}_B(A) \cdot \mathrm{p}(B), \text{ wenn } \mathrm{p}(B) \neq 0.$$



Wenn wir uns einmal $\mathrm{p}(A)$ als den "Anteil" von A-Ereignissen an der Gesamtmenge von Ereignissen vorstellen, dann gibt $p_B(A)$ den Anteil der $A \wedge B$-Ereignisse an den B-Ereignissen wieder. Um uns davon zu überzeugen, daß das eine gute Definition bedingter Wahrscheinlichkeit ist, wollen wir die Definition anhand eines Würfels veranschaulichen. Wie hoch ist die Wahrscheinlichkeit, daß eine Sechs gewürfelt wir? Wenn alles mit rechten Dingen

[26] Manche Autoren notieren die durch B bedingte Wahrscheinlichkeit von A als $\mathrm{p}(A|B)$. Auf den ersten Blick mag das so aussehen, als ob der Ausdruck $A|B$ Argument der Wahrscheinlichkeitsfunktion p ist – so wie es $A \wedge B$ oder $A \to B$ sein kann – und lädt die Suche nach einem Junktor $\circ$ ein, der bedingte Wahrscheinlichkeiten genau wiedergibt: $\mathrm{p}(A \circ B) = \mathrm{p}(A|B)$. Einen solchen Junktor kann es jedoch (normalerweise) gar nicht geben, wie wir gleich sehen werden; vgl. das Trivialitätsresultat von[183] auf pp. 257f.

[27] Weil Division durch 0 keinen Wert hat, ist die bedingte Wahrscheinlichkeit undefiniert im Fall $\mathrm{p}(B) = 0$. Man könnte hier durch eine Festsetzung nachhelfen – etwa so: wenn die Bedingung B unerfüllbar ist, dann sei $\mathrm{p}_B(A) = 1$. Aber das würde die merkwürdige Folge haben, daß bedingte Wahrscheinlichkeiten keine Wahrscheinlichkeiten sind. Denn wenn $\mathrm{p}(B) = 0$, dann $\mathrm{p}_B(A) = 1 = \mathrm{p}_B(\neg A)$, was im Widerspruch zu P3 stünde. Besser ist es, bedingte Wahrscheinlichkeit nur partiell, d.h. unter der Einschränkung zu definieren, daß die Bedingung erfüllbar sein muß.

zugeht, 1/6. Das ist 1/6 der Gesamtfläche im Bild rechts. Wie hoch ist die Wahrscheinlichkeit, daß eine Sechs gewürfelt wird (s), gegeben, daß eine gerade Zahl gewürfelt wird (g)? Um diese Frage zu beantworten betrachten wir nur die Felder unterhalb der Geraden und fragen, welchen Anteil daran die Fläche hat, auf die beide Eigenschaften s und g zutreffen. D.h. wir fragen nach dem Verhältnis von $s \wedge g$ und g. Die Wahrscheinlichkeit von $s \wedge g$ ist 1/6 (die gleiche wie die von s) und $\mathrm{p}(g) = 1/2$. Also ist das Verhältnis von $\mathrm{p}(s \wedge g)$ zu $\mathrm{p}(g)$ gleich $1/6 : 1/2 = 1/3$. Das beantwortet unsere Frage nach der Wahrscheinlichkeit von s gegeben die Bedingung g, d.h. nach dem Wert von $\mathrm{p}_g(s)$.

Die Definition von $\mathrm{p}_B(A)$ hat also genau den gewünschten Effekt. Sie schränkt die ursprüngliche Funktion p auf den Bereich ein, in dem B gilt und gibt die neue Wahrscheinlichkeit von A als "Anteil" von $A \wedge B$ an B an. Die Einschränkung auf die Bedingung B ist ferner minimal in dem Sinne, daß unter p und p_B alle Wahrscheinlichkeitsverhältnisse übereinstimmen, soweit sie nicht durch die Bedingung B (und dem, was daraus folgt) berührt sind. Denn

- für beliebige Formeln A und A' mit $A \models B$ und $A' \models B$ gilt:
$$\frac{\mathrm{p}_B(A)}{\mathrm{p}_B(A')} = \frac{\mathrm{p}(A)}{\mathrm{p}(A')}.$$

Schließlich können wir noch beobachten, daß die neue Funktion selbst wieder eine Wahrscheinlichkeitsfunktion ist:

- Wenn $\mathrm{p}(B) \neq 0$, dann ist p_B selbst eine Wahrscheinlichkeitsfunktion im Sinne von P0-2.

Es gilt deshalb insbesondere, daß die Funktion p_B die Gleichung P3 erfüllt, d.h.

P5. $$\mathrm{p}_B(\neg A) = 1 - \mathrm{p}_B(A).$$

Denn, nach (p_B) ist $\mathrm{p}_B(\neg A) = \mathrm{p}(\neg A \wedge B) : \mathrm{p}(B)$ und also ...

$$\begin{aligned}
&= (\mathrm{p}(\neg A \wedge B) + \mathrm{p}(A \wedge B) - \mathrm{p}(A \wedge B)) : \mathrm{p}(B) && \text{Arithmetik}\\
&= (\mathrm{p}(\neg A \wedge B) \vee (A \wedge B)) - \mathrm{p}(A \wedge B)) : \mathrm{p}(B) && \text{P2}\\
&= (\mathrm{p}(B) - \mathrm{p}(A \wedge B)) : \mathrm{p}(B) && \text{Logik, P0}\\
&= 1 - (\mathrm{p}(A \wedge B) : \mathrm{p}(B)) && \text{Arithmetik}\\
&= 1 - \mathrm{p}_A(B) && \text{Def. } \mathrm{p}_A.
\end{aligned}$$

Die Beobachtungen, die wir hier über Wahrscheinlichkeiten gemacht haben, gelten (beinahe) für beliebige Begriffe von Wahrscheinlichkeit, d.h. gleich ob wir Wahrscheinlichkeit etwa im Sinne von Häufigkeit, mit der sich ähnliche Ereignisse in der Welt wiederholen (oder, hypothetisch, wiederholen würden), oder im Sinne der Festigkeit von Überzeugungen interpretieren. Im folgenden wollen wir eine Beziehung herstellen zwischen dem Grad, in dem ein Sprecher eine Aussage behaupten darf und dem Grad seiner Überzeugung, daß die betreffende Aussage wahr sei. Für dieses Vorhaben ist Wahrscheinlichkeit als Maß der Festigkeit einer Überzeugung die einzig naheliegende.

Vier Hypothesen über indikativische Konditionale (MAT, BP, AT, SH). Man sollte nur das behaupten, wovon man genügend überzeugt ist; und wovon man genügend überzeugt ist, das kann man auch behaupten, wenn die Gelegenheit sich bietet. Das ist jedenfalls eine plausible Hypothese über den Zusammenhang von Behauptung und Überzeugung. Wenn wir den Grad der Überzeugung als Wahrscheinlichkeit wiedergeben, dann bedeutet das, daß die Behauptbarkeit b einer Aussage direkt mit ihrer Wahrscheinlichkeit variiert: [28]

BP. $$\mathrm{b}(A) = \mathrm{p}(A).$$

Nun haben wir zuvor festgestellt, daß in normalen Gesprächskontexten die Äußerung beliebiger Wahrheiten unangebracht ist. In diesem Sinne ist nicht jede Wahrheit jederzeit behauptbar ist. Was für Wahrheit gilt, gilt sicher auch für hohe Wahrscheinlichkeit. Nicht alles, was dem Sprecher genügend wahrscheinlich ist, kann er in beliebigen Situation angemessenerweise behaupten. Dagegegen stehen Erwägungen der Relevanz, der Schicklichkeit, des angemessenen Ausdrucks, der Ordnung der Rede etc. Aber, wenn wir von solchen Erwägungen einmal absehen, dann bleibt eine Klasse von Aussagen, die der Sprecher (im weiteren Sinne) behaupten kann, wenn – wie wir oben schon formuliert haben – sich eine passende Gelegenheit bietet. [29] Die Quantitätsmaxime als Kriterium für Behauptbarkeit in diesem Sinne ist zu streng, wie wir gesehen haben. Gibt BP ein erfolgreicheres Kriterium ab? Kann ein Specher genau das richtig behaupten (im weiteren Sinne), was er für hochwahrscheinlich hält?

[28] Die Behauptbarkeitsfunktion b möge die gleiche Metrik wie p haben, d.h. $\mathrm{b} : \mathrm{FML} \longrightarrow [0, 1]$.

[29] In der englischsprachigen Literatur wird in diesem Zusammenhang manchmal zwischen *assertibility* (mit *i*) und *assertability* (mit *a*) unterschieden. Letzteres ist ein Kunstwort, das den beschriebenen weiteren Sinn von Behauptbarkeit treffen soll.

Indikativische Konditionale wollen wir ab jetzt mit $A \succ B$ notieren. Für solche Konditionale bedeutet BP, daß

$$\mathrm{b}(A \succ B) = \mathrm{p}(A \succ B).$$

Wenn wir unter der Hypothese arbeiten, daß indikativische und materiale Konditionale die gleichen Wahrheitsbedingen haben, kurz

MAT. $$A \succ B \equiv A \rightarrow B,$$

dann folgt aus BP sogleich (per P0), daß

BPM. $$\mathrm{b}(A \succ B) = \mathrm{p}(A \rightarrow B).$$

Das ist aber keine gute Option. Denn nach (P2) wäre dann $\mathrm{b}(A \succ B) = \mathrm{p}(\neg A) + \mathrm{p}(B) - \mathrm{p}(\neg A \wedge B)$. Das erlaubt nun Instantiierungen nach diesem Muster: Man wähle A und B so, daß $\mathrm{p}(\neg A \wedge B)$ sehr klein obgleich $\mathrm{p}(\neg A)$ sehr groß ist. Dann wird $\mathrm{b}(A \succ B)$ ebenfalls groß sein.

> *Beispiel.* Es sei hochwahrscheinlich, daß Oma morgen zu Besuch kommt (A), und es sei sehr unwahrscheinlich, daß morgen ein Meteor die Erdkugel trifft und alles Leben auf diesem Planeten beendet (B). Dann stellt $\mathrm{p}(A \wedge B)$ nur einen sehr kleinen Wert (nahezu Null) dem hohen Wert von $\mathrm{p}(A) + \mathrm{p}(B)$ [$\approx \mathrm{p}(A)$] gegenüber. Aber dann ist $\neg A \succ B$, d.h. "Wenn Oma morgen nicht zu Besuch kommt, dann hat ein Meteor die Erde getroffen" im hohen Grade behauptbar.

Im Grunde wiederholt sich hier nur in einem probabilistischen Rahmen das Problem, auf das die Grice'sche Theorie antworten wollte. Die Unwahrscheinlichkeit des Antezedens eines Konditionals kann dieses unter bestimmten Umständen hochwahrscheinlich machen. Daraus sollten wir nicht auf eine hohe Behauptbarkeit des Konditionals schließen. BPM fordert aber genau dies und bringt uns daher keinen Schritt weiter.

Wenn wir BPM ablehnen, dann müssen wir entweder BP oder MAT aufgeben. Wir wollen hier zunächst weiter unter der Hypothese arbeiten, daß MAT die Wahrheitsbedingungen indikativischer Konditionale richtig erfaßt. Also müssen wir uns nun nach einem Ersatz für BP umsehen. Der Ersatz darf sich nicht zu weit vom Original entfernen, denn es kann ja nicht völlig falsch sein, daß Behauptbarkeit und Überzeugungsgrad in einem engen Verhältnis zueinander stehen.

Ein solcher Ersatz bietet sich schnell an. Schon recht deutlich formuliert, findet er sich in einer oft zitierten Fußnote zu einem Aufsatz von F.P. Ramsey (1929, [239, p. 247]):

> Wenn zwei Leute sich fragen, ob "Wenn A, dann B" der Fall ist und beide nicht wissen, ob A, dann fügen sie hypothetisch A ihrem Wissen hinzu und fragen auf dieser Basis, ob B. Gewissermaßen sind so "Wenn A, dann B" und "Wenn A, dann $\neg B$" kontradiktorisch. Wir können sagen, daß sie die Grade ihrer Überzeugungen, daß B, gegeben A, festlegen.

Diese Idee, den Status von $A \succ B$ in einem Überzeugungssystem zu prüfen, indem man den Status von B in einem leicht veränderten Überzeugungssystem prüft, nennt man den *Ramsey-Test.* In verschiedenen theoretischen Rahmen, nimmt der Test verschiedene Formen an. In der Theorie, in der wir uns jetzt bewegen, wollen wir ihn den *Adams Test* (oder Adams' These) nennen. Danach variiert die Behauptbarkeit eines Konditionals direkt mit der Stärke, mit welcher der Sprecher vom Konsequens überzeugt ist, wenn er die Wahrheit des Antezedens einmal hypothetisch voraussetzt:[30]

AT. $$\mathrm{b}(A \succ B) = \mathrm{p}_A(B).$$

Die Einschränkung der Wahrscheinlichkeit durch das Antezedens des zu prüfenden Konditionals übernimmt die Rolle der hypothetischen Annahme, von der Ramsey spricht. Das Resultat paßt auch gut zu Ramseys Beobachtung, daß die Behauptung von $A \succ B$ im Widerspruch zur Behauptung von $A \succ \neg B$ steht. Denn, wenn es sich so verhält, wie Adams es vorschlägt, dann ist $\mathrm{b}(A \succ \neg B) = 1 - \mathrm{p}_A(B)$ (nach P5) und also verhalten sich $\mathrm{b}(A \succ B)$ und $\mathrm{b}(A \succ \neg B)$ komplementär, d.h. $\mathrm{b}(A \succ B) = 1$ genau dann, wenn $\mathrm{b}(A \succ \neg B) = 0$.

Wir können uns leicht davon überzeugen, daß AT eine echte Alternative zu BPM ist. Denn $\mathrm{p}(A \to B)$ und $\mathrm{p}_A(B)$ können durchaus (sehr) verschiedene Werte annehmen. Zwar gilt (für $\mathrm{p}(A) \neq 0$)

$$\mathrm{p}_A(B) \leq \mathrm{p}(A \to B),$$

jedoch nicht die Umkehrung.

> *Beispiel.* Das Spiel bestehe aus 32 Karten. $\spadesuit 1$: Die erste Karte ist ein Pik; $\spadesuit 2$: die zweite Karte ist ein Pik. Dann ist $\mathrm{p}(\spadesuit 1) = 8/32 = 1/4$ und $\mathrm{p}_{\spadesuit 1}(\spadesuit 2) = 7/31$. Wir rechnen nun so:

[30] Vgl. Adams [2, 4].

$$
\begin{aligned}
\mathrm{p}(\spadesuit 1 \to \spadesuit 2) &= \mathrm{p}(\neg(\spadesuit 1 \wedge \neg\spadesuit 2)) && \text{P0}\\
&= 1 - (\mathrm{p}(\spadesuit 1) \cdot \mathrm{p}_{\spadesuit 1}(\neg\spadesuit 2)) && \text{P1, P4}\\
&= 1 - (\mathrm{p}(\spadesuit 1) \cdot (1 - \mathrm{p}_{\spadesuit 1}(\spadesuit 2))) && \text{P5}\\
&= 1 - (1/4 \cdot (1 - 7/31))\\
&= 1 - 24/124\\
&= 25/31
\end{aligned}
$$

$\mathrm{p}(\spadesuit 1 \to \spadesuit 2)$ ist also (bedeutend) größer als $\mathrm{p}_{\spadesuit 1}(\spadesuit 2)$.

Auch bei weiterer Betrachtung sieht AT gut aus. Die These bewährt sich in allen Fällen, in denen wir die einfache These MAT sichern wollen, indem wir Wahrheit und Behauptbarkeit auseinandertreten lassen. Das sind insbesondere solche Fälle, in denen $A \to B$ wahr ist, weil A falsch oder B wahr ist. Hier gilt es zu zeigen, daß die hohe Wahrscheinlichkeit von $\neg A$ oder von B nicht die hohe Wahrscheinlichkeit von $A \succ B$ garantiert. Wir benötigen also Beispiele mit $\mathrm{p}(\neg A) > \mathrm{p}_A(B)$ bzw. $\mathrm{p}(B) > \mathrm{p}_A(B)$.

Beispiele. Sei $\spadesuit \succ as$ das Konditional “Wenn die erste Karte ein Pik ist, dann ist sie ein As”. Die Wahrscheinlichkeit von $\neg\spadesuit$ ist recht hoch, nämlich $24/32 = 3/4$. Dagegen ist die Wahrscheinlichkeit des Konditionals, nämlich 1/8, deutlicher geringer. — Im Konditional “Wenn die erste Karte schwarz ist, dann ist sie kein Pik” hat das Konsequens die Wahrscheinlichkeit 3/4. Die entsprechende bedingte Wahrscheinlichkeit, welche die Behauptbarkeit des Konditionals mißt, beträgt aber nur 1/2.

Wie wenig wahrscheinlich ein Antezedens A auch sein mag, solange A einen positiven Wert hat, kann $\mathrm{p}_A(B)$ noch unwahrscheinlicher sein. (Ist $\mathrm{p}(A) = 0$, dann ist $\mathrm{p}_A(B)$ nicht definiert.) Ähnlich für ein hochwahrscheinliches Konsequens B: Solange $\mathrm{p}(B) < 1$, gibt es immer einen Wert ($\neq 0$) für A so, daß $\mathrm{p}_A(B) < \mathrm{p}(B)$. Damit ist AT in der Lage, die Fälle zu klären, die auch die Quantitätsmaxime zugunsten von MAT klären kann.

Die Quantitätsmaxime besagt auch: Wer A für falsch oder B für wahr hält, der kann das schwächere Konditional “Wenn A, dann B” nicht richtig behaupten! Das ist jedoch nicht immer so; vgl. p. 246. AT läßt die richtigen Ausnahmen von der Maxime zu: Wir können Wahrscheinlichkeiten vernünftig so verteilen, daß A niedrig oder B hoch ist (so, daß $\neg A$ bzw. B behauptbar sind), während $\mathrm{p}_A(B)$ hoch und also das Konditional $A \succ B$ behauptbar ist.

Beispiel. Peter glaubt fest an einen Sieg der Eintracht (B). Auf Meier kommt es diesmal nicht an. Ob dieser überhaupt spielen wird (A), steht auf Messers Schneide. Peter behauptet daher zurecht beides: Wenn Meier spielt, dann gewinnt die Eintracht, und: Wenn Meier nicht spielt, dann gewinnt die Eintracht. Es sei $\mathrm{p}(B) = 0,9$ und $\mathrm{p}(A) = 0,5$; dann ist $\mathrm{p}(A \wedge B) = 0,5 - \epsilon$, für einen kleinen Wert ϵ. Also hat $\mathrm{p}_A(B) = (0,5 - \epsilon) : 0,5$ ebenso wie $\mathrm{p}_{\neg A}(B)$ einen hohen Wert: etwa so hoch wie $\mathrm{p}(B)$.

Das Beispiel läßt sich leicht so verändern, daß auch mit einem sehr unwahrscheinlichen Antezedens A die Wahrscheinlichkeit von $\mathrm{p}_A(B)$ hoch genug ausfallen kann. (Man setze $\mathrm{p}(A) = 0,1$, d.h. Peter glaubt nicht, daß Meier spielen wird. Aber auf Meier kommt es diesmal ja nicht an ...)

Wenn wir AT folgen und die Behauptbarkeit eines Konditionals mit dessen Wahrscheinlichkeit messen, dann ist diese nicht einfach eine Funktion der Behauptbarkeiten des Antezedens und des Konsequens. Schon auf p. 249 haben wir gesehen, daß Junktoren keine Wahrscheinlichkeitsfunktionen sein müssen. Das gilt auch für Konditionale. Die Wahrscheinlichkeiten von A und A' mögen gleich sein und dennoch können die Konditionale $A \succ B$ und $A' \succ B$ unterschiedlich behauptbar sein.

Beispiel. Es sei $\spadesuit$: die erste Karte ist ein Pik, $\clubsuit$: die erste Karte ist ein Kreuz, und $\spadesuit as$: die erste Karte ist das Pik-As. Dann ist $\mathrm{p}(\spadesuit) = \mathrm{p}(\clubsuit) = 1/4$ während $\mathrm{p}_{\spadesuit}(\spadesuit as) = 1/8$ und $\mathrm{p}_{\clubsuit}(\spadesuit as) = 0$. Also ist $\mathrm{b}(\spadesuit \succ \spadesuit as) \neq \mathrm{b}(\clubsuit \succ \spadesuit as)$.

Das erlaubt uns, wie gewünscht, zwischen Paaren von Konditionalen wie (53) und (54) zu differenzieren:

(53) Wenn Meier ausfällt (A), gewinnt die Eintracht (B).
(54) Wenn der Torwart ausfällt (A'), gewinnt die Eintracht (B).

Beispiel. Peter ist vom Sieg der Eintracht überzeugt und hält den Ausfall des Torwarts für genauso unwahrscheinlich wie den von Meier. Gleichzeitig hält er die Mitwirkung des Torwarts für spielentscheidend, so daß er $\mathrm{p}_{A'}(B)$, im Gegensatz zu $\mathrm{p}_A(B)$, sehr niedrig ansetzt. Das erklärt, warum er (53), nicht aber (54) behauptet – obwohl doch nach der Quantitätsmaxime beide Konditionale gleich unbehauptbar wären, weil $\mathrm{p}(B)$ hoch und im Falle von (53) darüberhinaus noch $\mathrm{p}(A)$ niedrig ist.

Adams' These ist also nicht nur vorab sehr plausibel, sondern besteht auch erfolgreich einige Proben auf Schlüsseleigenschaften von Konditionalen. Aber wie kommt es zu dieser erfolgreichen Ausnahme von BP? Wie können wir erklären, warum AT die richtige Beziehung zwischen Behauptbarkeit und Wahrscheinlichkeit im Falle von Konditionalen herstellt?

Eine einfache Erklärung würde *Stalnakers Hypothese* liefern:

SH. $$\mathrm{p}(A \succ B) = \mathrm{p}_A(B).$$

Danach ist die Wahrscheinlichkeit eines indikativischen Konditionals nichts anderes als die entsprechende konditionale Wahrscheinlichkeit. Diese These hat den Vorzug der Einfachheit und erlaubt, AT direkt aus BP herzuleiten; keine Ausnahme für Konditionale also:

$$\dfrac{\dfrac{}{\mathrm{b}(A \succ B) = \mathrm{p}(A \succ B)}\,\mathrm{BP} \qquad \dfrac{}{\mathrm{p}(A \succ B) = \mathrm{p}_A(B)}\,\mathrm{SH}}{\mathrm{b}(A \succ B) = \mathrm{p}_A B}$$

SH ist eine These mit Biss. Denn falls sie richtig sein sollte, dann ist die einfache These MAT über die Wahrheitsbedingungen von indikativischen Konditionalen falsch. Aus SH und MAT könnten wir aufgrund von P0 die falsche Gleichung $\mathrm{p}(A \to B) = \mathrm{p}_A(B)$ ableiten. Wir müssen uns also zwischen SH und MAT entscheiden. Im nächsten Abschnitt erklären wir, warum SH für MAT keine Gefahr darstellen kann.

Das Trivialitätsresultat von Lewis. Wir wissen bereits, daß die materiale Implikation $\to$ kein Junktor von der Art ist, daß für beliebige Sätze A und B und beliebige Wahrscheinlichkeitsfunktionen p, die Wahrscheinlichkeit von $A \to B$ durch die entsprechende konditionale Wahrscheinlichkeit wiedergegeben werden kann. Im allgemeinen gilt eben nicht, daß $\mathrm{p}(A \to B) = \mathrm{p}_A(B)$. Woher wissen wir eigentlich, daß es überhaupt Junktoren mit dieser Eigenschaft gibt? Stalnakers Hypothese SH setzt genau das offensichtlich voraus: daß es nämlich einen zweistelligen Junktor $\circ$ gibt, so daß für beliebige Sätze A und B und beliebige Wahrscheinlichkeitsfunktionen p, die Wahrscheinlichkeit von $A \circ B$ mit der entsprechenden konditionalen Wahrscheinlichkeit zusammenfällt, d.h.

($\circ$) $$\mathrm{p}(A \circ B) = \mathrm{p}_A(B).$$

In diesem Abschnitt werden wir zeigen, daß diese Voraussetzung normalerweise falsch ist. Das ist das sogenannte Trivialitätsresultat von Lewis [183]. Die Einschränkung “normalerweise” soll triviale Wahrscheinlichkeitsfunktionen ausschließen, die zwei Sätzen A und (kontingentem) B keine Werte zuordnen können, ohne daß die Wahrscheinlichkeit von B mit der von $A \circ B$ zusammenfällt.[31]

[31] Das Trivialitätsresultat von Lewis gibt es in verschieden starken Varianten. Zwei Varianten finden sich in [183] und [186]; weitere in [121] und [122].

Satz 7. (Trivialitätsresultat von Lewis) *Wenn* $\mathrm{p}(A \circ B) = \mathrm{p}_A(B)$ *(für* $\mathrm{p}(A) > 0$*) und* $0 < \mathrm{p}(B) < 1$*, dann* $\mathrm{p}(A \circ B) = \mathrm{p}(B)$*.*

Zunächst beweisen wir ein Lemma.

Lemma 8. *Jeder Junktor* $\circ$*, der die Gleichung* $(\circ)$ *erfüllt, hat auch die Eigenschaft* $\mathrm{p}_A(B \circ C) = \mathrm{p}_{A\wedge B}(C)$*.*

Beweis.

$$\begin{aligned}
\mathrm{p}_A(B \circ C) &= (\mathrm{p}_A)_B(C) && (\circ)\\
&= \frac{\mathrm{p}_A(B \wedge C)}{\mathrm{p}_A(B)} && \text{Def. } (\mathrm{p}_A)_B\\
&= \frac{\mathrm{p}(A \wedge B \wedge C) : \mathrm{p}(A)}{\mathrm{p}(A \wedge B) : \mathrm{p}(A)} && \text{Def. } \mathrm{p}_A\\
&= \frac{\mathrm{p}(A \wedge B \wedge C)}{\mathrm{p}(A \wedge B)} && : \mathrm{p}(A)\\
&= \mathrm{p}_{A\wedge B}(C) && \text{Def. } \mathrm{p}_{A\wedge B}
\end{aligned}$$

∎

Den Trivialitätssatz beweisen wir nun, indem wir von der folgenden Eigenschaft einer Wahrscheinlichkeitsfunktion p ausgehen. Für eine kontingente Aussage B (d.h. $0 < \mathrm{p}(B) < 1$) gilt:

(1) $$\mathrm{p}(A) = \mathrm{p}_B(A) \cdot \mathrm{p}(B) + \mathrm{p}_{\neg B}(A) \cdot \mathrm{p}(\neg B).$$

Denn $A \equiv (A \wedge B) \vee (A \wedge \neg B)$ und also $\mathrm{p}(A) = \mathrm{p}((A \wedge B) \vee (A \wedge \neg B))$ (nach P0). Da $A \wedge B$ und $A \wedge \neg B$ einander ausschließen, ist (nach P2)

$$\mathrm{p}(A) = \mathrm{p}(A \wedge B) + \mathrm{p}(A \wedge \neg B).$$

Daraus folgt (1) nach P4 und der Annahme $0 < \mathrm{p}(B) < 1$.

Eine Instanz von (1) ist

(2) $$\mathrm{p}(A \circ B) = \mathrm{p}_B(A \circ B) \cdot \mathrm{p}(B) + \mathrm{p}_{\neg B}(A \circ B) \cdot \mathrm{p}(\neg B).$$

Jetzt wenden wir das Lemma auf (2) an und erhalten

(3) $$\mathrm{p}(A \circ B) = \mathrm{p}_{A\wedge B}(B) \cdot \mathrm{p}(B) + \mathrm{p}_{A\wedge\neg B}(B) \cdot \mathrm{p}(\neg B).$$

Nun ist aber

(4) $$\mathrm{p}_{A\wedge B}(B) = 1 \quad \text{und} \quad \mathrm{p}_{A\wedge\neg B}(B) = 0.$$

Setzen wir diese Werte in (3) ein, so erhalten wir

$$\mathrm{p}(A \circ B) = \mathrm{p}(B).$$

∎

Nun wissen wir, daß Stalnakers Hypothese SH über indikativische Konditionale $A \succ B$ nur in trivialen Fällen zutreffen kann. Damit ist die einfachste Erklärung von Adams' These $\mathrm{b}(A \succ B) = \mathrm{p}_A(B)$ vom Tisch: Die These AT ist nicht einfach nur ein Fall des allgemein plausiblen Prinzips BP, wonach Behauptbarkeit durch Wahrscheinlichkeit gemessen wird. Die These ist dem allgemeinen Prinzip zwar nahe genug, um von dessen Plausibilität zu profitieren – aber sie stellt doch eine Ausnahme von dem Prinzip dar. Es wäre gut, wenn wir dafür eine Erklärung hätten.

8. Konditionale und Robustheit

Wir wissen, daß $\mathrm{p}_A(B) \leq \mathrm{p}(A \to B)$ (wenn $\mathrm{p}(A) \neq 0$) und also, daß es für alle A und B einen Wert $0 \leq x \leq 1$ gibt, so daß

$$(*) \qquad \mathrm{p}_A(B) = \mathrm{p}(A \to B) - x.$$

Wenn nun, nach AT, $\mathrm{p}_A(B)$ die Behauptbarkeit eines Konditionals $A \succ B$ mißt, dann folgt aus $(*)$, daß diese durch die Wahrscheinlichkeit des materialen Konditionals $A \to B$ mit einem gewissen Abzug gegeben ist; d.h. AT behauptet dies:

$$\text{AT.} \qquad \exists x \ \mathrm{b}(A \succ B) = \mathrm{p}(A \to B) - x.$$

In jeder Theorie also, die x (abhängig von A und B) so bestimmt, daß die Gleichung $(*)$ erfüllt ist, wäre AT ableitbar. Und wenn x so bestimmt wird, daß deutlich ist, warum der Wert von x die Behauptbarkeit des Konditionals steuert, dann liegt eine Ableitung von AT mit erklärender Kraft vor: eine Ableitung, die erklärt, warum $\mathrm{p}_A(B)$ die Behauptbarkeit des entsprechenden Konditionals wiedergibt.

Eine erste solche Theorie hat Lewis [183] vorgelegt, sie aber schon bald verworfen, um sich der Theorie von Jackson anzuschließen. Der Grundgedanke von Jacksons [146, 148, 149] Theorie beruht auf der Beobachtung, daß die Äußerung von Konditionalen die Funktion hat, zu einem Modus Ponens-Schritt einzuladen. Ein Konditional $A \succ B$ ist gewissermaßen ein Fahrschein zum Schließen: Wenn A eintrifft, dann berechtigt der Fahrschein – das Konditional –, zu B überzugehen.[32] So verstehen Hörer jedenfalls einen Sprecher, der ein Konditional äußert. Deshalb ist es grob irreführend, ein Konditional zu äußern, bloß weil man das Antezedens für sehr unwahrscheinlich hält. In diesem Fall würde der Sprecher ja nicht, wie scheinbar angekündigt, schließen, sondern das Konditional zurückziehen, sobald

[32] Diese Metapher geht auf Gilbert Ryle zurück, der Konditionale als "inference tickets" bezeichnet hat.

er erfährt, daß das Antezedens der Fall ist. Im Bild: Statt zum Übergang von A nach B zu berechtigen, würde der Fahrschein $A \succ B$ in dem Moment seine Gültigkeit verlieren, in dem A eintrifft. Oder, wie Jackson sagt: Das Konditional wäre hinsichtlich seines Antezedens nicht *robust.*

Beispiel. Peter glaubt (mit hoher Wahrscheinlichkeit), daß Aigner das Tor geschossen hat. Daraufhin behauptet er: "Wenn Aigner das Tor nicht geschossen hat, dann war es Meier". Paul überzeugt nun Peter, daß Aigner nicht geschossen hat und sagt weiter: "Nach dem, was Du gesagt hast, Peter, glaubst Du also, daß es Meier war." – "Nein, das glaube ich nicht", antwortet Peter. — ??

Eine Aussage A ist robust gegenüber einer Aussage B (hinsichtlich einer Wahrscheinlichkeitsfunktion p), wenn die Wahrscheinlichkeit von A unter der Annahme B nicht leidet (kurz: $\mathrm{p}(A) \approx \mathrm{p}_B(A)$).[33] Konditionale sind dann und nur dann behauptbar, wenn sie genügend hoch wahrscheinlich und robust gegenüber ihrem Antezedens sind, d.h. wenn sie nicht an Wahrscheinlichkeit verlieren, sobald das Antezedens an Wahrscheinlichkeit gewinnt. Nur so kann ein Konditional seine Funktion als Folgerungsfahrschein erfüllen. Ein Konditional ist also in dem Grade behauptbar, in dem wir es inbesondere auch dann akzeptieren, wenn sein Antezedens sich als wahr erweist. Diese Grundidee können wir so zusammenfassen:

(†) $A \succ B$ ist genau dann behauptbar, wenn $\mathrm{p}(A \succ B)$ und $\mathrm{p}_A(A \succ B)$ hoch sind.

Wenn wir nun die Äquivalenzthese MAT annehmen, dann können wir $\rightarrow$ für $\succ$ einsetzen und erhalten (h für "hoch"):

(‡) $\mathrm{b}(A \succ B) \approx h$ gdw $\mathrm{p}(A \rightarrow B) \approx h$ und $\mathrm{p}_A(A \rightarrow B) \approx h$.

Da $\mathrm{p}_A(A \rightarrow B) \leq \mathrm{p}(A \rightarrow B)$, so impliziert ein hoher Wert des linken Ausdrucks einen mindestens so hohen Wert des rechten Ausdrucks. Aus (‡) folgt daher

JT. $$\mathrm{b}(A \succ B) \approx h \quad \text{gdw} \quad \mathrm{p}_A(A \rightarrow B) \approx h.$$

Das ist *Jacksons These.* Sie beruht auf der These MAT und der Berücksichtigung der Robustheitsforderung für behauptbare Konditionale. Jackson nennt sie daher das Kernstück einer "ergänzten Äquivalenztheorie". Aus JT

[33] Das ist ein etwas stärkerer Begriff von Robustheit, als er für Jacksons These (s.u.) benötigt wird. Für diese These würde es genügen zu sagen, daß A gegenüber B robust ist, falls die *hohe* Wahrscheinlichkeit von B unter der Annahme A nicht zurückgeht. Für die Herleitung von AT in voller Allgemeinheit, werden wir jedoch auf die stärkere Definition zurückgreifen müssen.

läßt sich schnell AT für Konditionale mit hoher Behauptbarkeit herleiten. Denn

$$\begin{aligned}
\mathrm{p}_A(A \to B) &= \mathrm{p}_A(\neg A \vee B) && \text{P0} \\
&= \frac{\mathrm{p}((\neg A \vee B) \wedge A)}{\mathrm{p}(A)} && \text{Def. } \mathrm{p}_A \\
&= \frac{\mathrm{p}(A \wedge B)}{\mathrm{p}(A)} && \text{P0} \\
&= \mathrm{p}_A(B) && \text{Def. } \mathrm{p}_A
\end{aligned}$$

Also besagt JT, daß $A \succ B$ genau dann eine hohe Behauptbarkeit hat, wenn $\mathrm{p}_A(B)$ hoch ist. So kann JT erklären, warum AT erfolgreich Voraussagen über die Behauptbarkeit von Konditionalen macht. (Da diese Erklärung von MAT Gebrauch macht, möchte Jackson [151, p. 51] sie als indirekten Beleg für die Richtigkeit der Äquivalenzthese verstehen. Aber genau besehen, genügt es für die Herleitung beim Übergang von (†) zu (‡) irgendeinen Junktor $\circ$ für $\succ$ mit der Eigenschaft $A \wedge B \equiv (A \circ B) \wedge A$ (Zeile 3 der Ableitung) einzusetzen; vgl. dazu Ellis [72].)

Soweit haben wir nur hohe Wahrscheinlichkeiten betrachtet. Aber Adams' These AT stellt auch für kleine Behauptbarkeits- und Wahrscheinlichkeitswerte eine Korrelation her. Um Adams' These ganz allgemein herzuleiten, erinnern wir an die Definition von Robustheit:

B ist genau dann robust gegenüber A, wenn $\mathrm{p}(B) \approx \mathrm{p}_A(B)$.

Die Differenz $\mathrm{p}(B) - \mathrm{p}_A(B)$ gibt den Grad der Verletzlichkeit ("Irrobustheit") von B gegenüber der Annahme A an. Es stellt sich heraus, daß genau dieser Verletzlichkeitsgrad die Rolle des Subtrahenden in der Gleichung $(*)$ übernehmen kann. Im Falle eines Konditionals $A \to B$ ist die Verletzlichkeit des Konditionals gegenüber der Annahme des Antezedens gegeben durch

$$x \qquad \mathrm{p}(A \to B) - \mathrm{p}_A(A \to B).$$

Setzen wir diesen Ausdruck für x ein in

$$\mathrm{b}(A \succ B) = \mathrm{p}(A \to B) - x$$

(siehe (*) oben), so erhalten wir

$$\begin{aligned}
\mathrm{b}(A \succ B) &= \mathrm{p}((A \to B) - [\,\mathrm{p}(A \to B) - \mathrm{p}_A(A \to B)\,] \\
&= \mathrm{p}((A \to B) - [\,\mathrm{p}(A \to B) - \mathrm{p}_A(B)\,] && \text{s.o.} \\
&= \mathrm{p}_A(B) && \text{Arithmetik}
\end{aligned}$$

Die Behauptbarkeit des indikativischen Konditionals nähert sich also in dem Maße der Wahrscheinlichkeit des entsprechenden materialen Konditionals, in dem letzteres robust gegenüber seinem Antezedens ist.

Man beachte, daß nur die Robustheit gegenüber der Wahrheit des Antezedens, nicht gegenüber der Falschheit des Konsequens eine Rolle für die Behauptbarkeit des Konditionals spielt. So wird die Möglichkeit gelassen, aus der hohen Behauptbarkeit von $A \succ B$ und $\neg B$ nicht auf die hohe Behauptbarkeit von $\neg A$ zu schließen. (Der Folgerungsfahrschein gilt nur für die Modus Ponens-, nicht für die Modus Tollens-Richtung.)

> *Beispiel*: Aus "Wenn Peter nicht in Frankfurt wohnt, dann wohnt er anderswo im Rhein-Main-Gebiet" und "Peter wohnt nicht im Rhein-Main-Gebiet" wollen wir nicht auf "Peter wohnt in Frankfurt" schließen.

Nun wird auch deutlich, warum AT nicht einfach das Schema BP instanziiert. Denn bei genauerer Betrachtung erkennen wir, warum wir uns von BP (in der Richtung von rechts nach links) verabschieden müssen. Viele Aussagen drücken nicht nur aus, daß ihre Wahrheitsbedingung erfüllt ist, sondern signalisieren darüber hinaus robuste Wahrscheinlichkeit oder andere Implikaturen (s.u.). Der richtige Umgang mit solchen Implikaturen gehört zum kompetenten Gebrauch einer ganzen Reihe sprachlicher Ausdrücke. In all diesen Fällen gilt, daß hohe Wahrscheinlichkeit eine notwendige Bedingung für Behauptbarkeit ist – jedoch nicht umgekehrt. Zur hohen Wahrscheinlichkeit von A muß eine weitere, durch die Implikatur angezeigte Bedingung hinzutreten, damit A behauptbar ist. Im Falle von Konditionalen ist das die Bedingung der Robustheit gegenüber dem Antezedens.

Daß es sich hierbei um ein sehr allgemeines Phänomen handelt, verdeutlichen wir uns am besten an einem Beispiel, das mit Robustheit nichts zu tun hat. So ist der Satz A*, obwohl* B genau dann hochwahrscheinlich, wenn die Konjunktion $A \wedge B$ eine hohe Wahrscheinlichkeit hat. Aber eine hohe Wahrscheinlichkeit von $A \wedge B$ reicht nicht aus, den Satz A*, obwohl* B behauptbar zu machen. Hier wird ein Kontrast ausgedrückt, der in der einfachen Konjunktion fehlt. Zwar sind $A \wedge B$ und A*, obwohl* B unter denselben Bedingungen wahr, nämlich wenn A und B wahr sind, aber die Bedeutung von A*, obwohl* B wird von dieser Wahrheitsbedingung nicht vollständig ausgeschöpft. Vielmehr trägt A*, obwohl* B eine *Bedeutungsimplikatur*, nämlich, daß die gleichzeitige Wahrheit von A und B überraschend ist. Grice nennt solche Bedeutungsimplikaturen auch *konventionelle Implikaturen*, um sie von Konversationsimplikaturen abzugrenzen. Im Falle von A*, obwohl* B können wir daher nicht einfach

$$\mathrm{b}(A,\ \textit{obwohl}\ B) = \mathrm{P}(A,\ \textit{obwohl}\ B)$$

setzen, denn die rechte Seite ist (aufgrund von P0) gleich $\mathrm{p}(A \wedge B)$. Vielmehr ist *A, obwohl B* nur dann behauptbar, wenn die an "obwohl" haftende Bedeutungsimplikatur eines Kontrastes zutrifft.

Robustheit gehört ebenfalls zu den Bedeutungsimplikaturen und betrifft nicht nur Konditionale. Zum Beispiel sollte niemand die Disjunktion

(55) Peter ist ein Fanatiker (A), oder jedenfalls sehr verblendet (B),

behaupten, der Peter weder für einen Fanatiker noch für in anderer Weise verblendet hält. Behauptbarkeit impliziert also hohe Wahrscheinlichkeit. Umgekehrt muß aber für die Behauptbarkeit von (55) zur einfachen Disjunktion $A \vee B$ noch die Robustheit gegenüber $\neg A$ hinzutreten, d.h. der Sprecher sollte der Aussage auch dann noch eine hohe Wahrscheinlichkeit zumessen, wenn er A für falsch hält, $\mathrm{p}_{\neg A}(A \vee B)$ also hoch ist. Es ist nicht erstaunlich, daß Sprachen über Konstruktionen verfügen, die Robustheit signalisieren. Oft wissen wir nicht, was unsere Gesprächspartner glauben, wollen aber, daß manche unserer Aussagen auch dann auf Gehör und Glauben treffen, wenn der Gesprächspartner mit anderen Aussagen nicht einverstanden ist. Genau deshalb äußern wir den schwächeren Satz (55), obwohl wir eigentlich die stärkere Aussage A glauben: Wir wollen etwas behaupten, daß robust ist im Austausch mit Gesprächspartnern, die A für falsch halten.

Bedeutungsimplikaturen sind im Unterschied zu den aus allgemeinen Maximen abzuleitenden Gesprächsimplikaturen sehr spezifische Konventionen: Während Gesprächsimplikaturen Folge allgemeiner pragmatischer Regeln zur effizienten Gesprächsführung sind, betreffen Bedeutungsimplikaturen bestimmte Ausdrücke und finden ihren Niederschlag in Lexikoneinträgen. Sprecher erlernen solche Implikaturen als Teile der Bedeutung von Ausdrücken. Wie Gesprächsimplikaturen – und im Unterschied zu Implikationen aus den Wahrheitsbedingungen – können Bedeutungsimplikaturen ohne logischen Widerspruch gestrichen, d.h. verneint werden. Dem Satz

(56) Die Eintracht hat gewonnen, *obwohl* sie schlecht aufgestellt war

kann Peter ohne *logischen* Widerspruch den Satz

(57) Es überrascht mich überhaupt nicht, daß eine schlecht aufgestellte Mannschaft gewinnt

hinzufügen. Die Wahrheit von (56) schließt (57) nicht aus. Aber indem Peter (57) äußert, gibt er zu erkennen, daß er die Bedeutung von "obwohl", d.h. die Implikatur eines Kontrastes, nicht erfaßt hat. Deshalb untergräbt die Äußerung von (57) zwar nicht die Wahrheit, wohl aber die Behauptbarkeit von (56). Das ist, wie wir gesehen haben, bei Gesprächsimplikaturen anders.

Sie zu streichen, untergräbt weder die Wahrheit noch die Behauptbarkeit der betreffenden Äußerung.

Nach Jackson gehört es nicht zur Bedeutung *im engeren Sinne*, d.h. zur Wahrheitsbedingung eines indikativischen Konditionals, daß es robust im Hinblick auf sein Antezedens ist. Seine Bedeutung in diesem engeren Sinne ist die des materialen Konditionals. Aber die Robustheit gehört wesentlich zur Konvention des richtigen Gebrauchs eines indikativischen Konditionals. In diesem Sinne ist die Robustheit gegenüber dem Antezedens Teil der Bedeutung *im weiteren Sinne* eines Konditionals. Ähnlich wie im Falle von *obwohl* verhält es sich hier so: Wenn Peter nicht glaubt, daß Meier das Tor geschossen hat, obgleich er zuvor das Konditional

Wenn Aigner das Tor nicht geschossen hat, dann war es Meier

behauptet und später erfahren hat, daß Aigner nicht der Torschütze war, dann sind Peters Überzeugungen zwar nicht inkonsistent – aber sie geben zu erkennen, daß er einen wichtigen Teil der Bedeutung von *wenn-dann* nicht erfaßt hat. Er hat nicht verstanden, daß er mit der Äußerung des Konditionals eine bestimmte Art von Robustheit signalisiert hat.

9. Probabilistisches Folgern mit Konditionalen

Sehen wir uns Lewis' Trivialitätsresultat gegen Stalnakers Hypothese

SH. $$\mathrm{p}(A \succ B) = \mathrm{p}_A(B)$$

genauer an, dann entdecken wir im Beweis eine Annahme, die in der Formulierung des Satzes nicht explizit gemacht wurde. Das Lemma zum Resultat beginnt so:

(∗) $$\mathrm{p}_A(B \succ C) = (\mathrm{p}_A)_B(C),$$

und diese Gleichung wird als Instanz von SH deklariert. Man beachte, daß die rechte Seite der Gleichung nach SH wiederum gleich ist zu

$$\mathrm{p}(A \succ (B \succ C)).$$

Die nicht explizite Annahme ist also, daß indikativische Konditionale (links) iterierbar sind. Die Annahme der Iterierbarkeit ist zwingend, wenn wir davon ausgehen, daß solche Konditionale wahrheitsdefinit sind: Wenn A und B einen Wahrheitswert haben, dann hat auch $A \succ B$ einen Wahrheitswert. Ein einfaches induktives Argument zwingt dann zur Konklusion, daß beliebige eingebettete und also auch iterierte Konditionale wahrheitsdefinit sind – alles andere wäre *ad hoc*.

Dem Trivialitätsresultat von Lewis können wir entgehen, indem wir die Iterierbarkeit von Konditionalen ablehnen, was wiederum erfordert, die These von der Wahrheitsdefinitheit indikativischer Konditionale – und also inbesondere auch MAT – abzulehnen. Wir können das Argument gegen SH dann so umdrehen:

Statt: Aus dem Trivialitätsresultat und der Wahrheitswertigkeit von Konditionalen schließen wir auf die Falschheit von SH.

Nun: Aus dem Trivialitätsresultat und SH – richtig verstanden! – schliessen wir darauf, daß indikativische Konditionale nicht wahrheitswertig sind.

Richtig verstehen wir SH, wenn wir dessen rechte Seite, $\mathrm{p}_A(B)$, als Erläuterung der problematischen linken Seite, $\mathrm{p}(A \succ B)$, auffassen. Problematisch ist der Ausdruck $\mathrm{p}(A \succ B)$, weil die Funktion p eigentlich den Grad einer Überzeugung, daß etwas der Fall ist, mißt. Aber wir wollen ja jetzt Abstand nehmen von der Annahme, daß $A \succ B$ irgendeine Überzeugung darüber, wie die Welt ist, beschreibt. So kann $A \succ B$ nur in einem uneigentlichen, zu erläuternden Sinne in die Argumentstelle von p geraten. Die rechte Seite von SH liefert die Erläuterung. Der Ausdruck "die Wahrscheinlichkeit von $A \succ B$" bedeutet nichts anderes als die Wahrscheinlichkeit von B unter der Annahme A, wobei A und B Aussagen sind, die im eigentlichen Sinne Argumente von p, d.h. wahr oder falsch ("faktisch") sein können. Das ist Adams' Hypothese, der Kern seiner "probability conditional theory",[34]

AH. $$\mathrm{p}(A \succ B) = \begin{cases} \mathrm{p}_A(B), & \text{falls } \mathrm{p}(A) > 0; \\ 1 & \text{anderenfalls.} \end{cases}$$

Faktisch wollen wir einen Satz A in einer Sprache mit Booleschen Verknüpfungen und dem Junktor $\succ$ nennen, falls A ein Atom ist oder allein mit Booleschen Verknüpfungen, also rein wahrheitsfunktional aufgebaut ist.[35] Im Verein mit dem, was wir Adams' These AT genannt haben, erhalten wir so eine Behauptbarkeitsbedingung für indikativische Konditionale, ohne daß wir annehmen müssen, daß solche Konditionale darüberhinaus auch Wahrheitsbedingungen haben; dieser weiteren Annahme steht Lewis' Resultat entgegen, wie wir gesehen haben. Nach Adams behauptet man mit $A \succ B$ zwar konditional, daß B wahr, nicht jedoch, daß das Konditional

[34] Die Festsetzung für den Fall $\mathrm{p}(A) = 0$ hat in Adams' Schriften einen etwas unsicheren Status. Manchmal ist sie explizit angegeben, manchmal scheint Adams die Option anzudeuten, daß in diesem Fall die Wahrscheinlichkeit des Konditionals nicht definiert ist. Jedenfalls spielt dieser Fall keine kritische Rolle in Adams' Theorie.

[35] Das ist die einfache Definition. Es liegt jedoch nichts daran, daß die Verknüpfungen wahrheitsfunktional sind. Sie müssen nur, im Gegensatz zu $\succ$, Sätze erzeugen, deren semantische Werte Gegenstände von Überzeugungen sein können.

wahr ist. Quine (1950) bringt das so auf den Punkt (ohne weiter etwas daraus zu machen):

> Eine Behauptung der Form *Wenn A, dann B* gilt gewöhnlich weniger als Behauptung eines Kondtionals, denn als konditionale Behauptung des Konsequens. [234, §3]

Das ist die Funktion indikativischer Konditionale in der Rede: Sie sind Mittel zur hypothetischen Behauptung einer Wahrheit, ohne selbst eine Wahrheit kategorisch behaupten zu wollen. [36]

> *Bemerkung.* Die Formulierung AH mag nicht ganz glücklich sein, denn eine Wahrscheinlichkeitsfunktion p kann ja – wie wir gerade gesagt haben – nur in einem "uneigentlichen" Sinne auf indikativische Konditionale angewandt werden. Formulierungen wie "Konditionale sind nicht wahr oder falsch, sondern nur wahrscheinlich" oder "die Wahrscheinlichkeit von Konditionalen ist nicht die Wahrscheinlichkeit ihrer Wahrheit" decken die Schwierigkeit nur mehr oder weniger elegant zu. Der Vorschlag, $\mathrm{p}(A \succ B)$ sei nur eine Ersatznotation für $\mathrm{p}_A(B)$, mag technisch genügen, weist aber AH nicht mehr als den Status einer unverbindlichen Vereinbarung zu. Besser wäre es, auf AH als Trittstein zu verzichten und direkt mit AT zu arbeiten, d.h. einer These nicht über die Wahrscheinlichkeit, sondern über die *Behauptbarkeit* von Konditionalen. Vieles, was Adams schreibt, scheint genau das nahezulegen. Wie dem auch sei, wir werden hier Adams' Entscheidung folgen, die Theorie allein mit Hilfe von Wahrscheinlichkeitsfunktionen darzustellen.

* * *

Diese Theorie steht zunächst einer prinzipiellen Schwierigkeit gegenüber. Ein gültiger Schluß ist einer, der die Wahrheit der Prämissen auf die Konklusion ausnahmslos überträgt. Wenn indikativische Konditionale weder wahr noch falsch sind, wie können sie dann wesentliche Bestandteile von Argumenten sein, die wir im Hinblick auf ihre Gültigkeit beurteilen? Es scheint, als könnten indikativische Konditionale gar keine Logik haben. Aber das ist sicher absurd. Es sei nur an die Funktion von solchen Konditionalen als Folgerungsfahrscheine erinnert; d.h. wir fordern, daß Modus Ponens gültig

[36] Wir haben hier Lewis' Trivialitätsresultat als prinzipiellen Beweggrund für die Ablehnung von Wahrheitsbedingungen für indikativische Konditionale in den Vordergrund gestellt. Viele Autoren führen dafür aber weitere und von Lewis' Resultat unabhängige Gründe an; vgl. z.B. Gibbard [108], Appiah [14], Edgington [69], [71], Bennett [25], Arló-Costa [17] und Adams [4].

sei: Aus $A \succ B$ und A folgt B. Aber in welchem Sinne kann dieser Schluß gültig sein, wenn $A \succ B$ gar nicht wahrheitswertig ist?

Wir benötigen einen Begriff gültiger Folgerung, der auch auf Prämissen und Konklusionen anwendbar ist, von denen wir nur annehmen, daß sie im Hinblick auf ihre Wahrscheinlichkeit – und möglicherweise nicht im Hinblick auf ihre Wahrheit – bewertet werden können. Dieser sollte weiter sein als der übliche Begriff logischer Folgerung, ihn also als Grenzfall enthalten. Auf dem Weg zu einer solchen Relation gültiger Folgerung, kommen wir auf die Beobachtung,

PL. $\quad$ wenn $A \models B$, dann $(\forall \mathrm{p})\ \mathrm{p}(A) \leq \mathrm{p}(B)$

zurück, vgl. p. 249. PL nennt eine Eigenschaft von $\models$, die wir uns insbesondere dann wünschen, wenn wir aus unsicheren Prämissen schließen: Wenn unter jeder Wahrscheinlichkeitsverteilung p der Wert von B nie kleiner als der von A werden kann, dann garantiert der Schluß von A nach B, daß es unter keiner Funktion p zu einem Abfall der Wahrscheinlichkeit von der Prämisse zur Konklusion kommt. Auch die Umkehrung von PL gilt, wenn A und B wahrheitswertige Aussagen sind. (Denn Wahrheitswertverteilungen $I : \mathrm{FML} \longrightarrow \{0, 1\}$ sind Grenzfälle von Wahrscheinlichkeitsverteilungen. Also folgt aus der Annahme $\mathrm{p}(A) \leq \mathrm{p}(B)$ $(\forall \mathrm{p})$, daß dies insbesondere auch für alle I gilt, falls A und B als Argumente für I zulässig sind. Da nun $A \models B$ gdw $\forall I : I(A) \leq I(B)$, so folgt $A \models B$.)

Diese Beobachtung gibt so etwas wie eine wahrscheinlichkeitstheoretische Interpretation logischer Folgerung (aus einer Prämisse) ab. Für die weitere Entwicklung dieses Gedankens wird es einfacher sein mit dieser Formulierung zu arbeiten: Gültige Schlüsse (aus einer Prämisse) haben die Eigenschaft, daß die Konklusion nie, d.h. unter keiner Wahrscheinlichkeitsverteilung, *ungewisser* als die Prämisse sein kann. Ungewißheit ist hier einfach Unwahrscheinlichkeit (das Komplement von Wahrscheinlichkeit), d.h. ein Maß $\overline{\mathrm{p}}$ so, daß $\overline{\mathrm{p}}(A) = 1 - \mathrm{p}(A)$. Wenn wir – wie dies meist der Fall ist – von ungewissen Annahmen ausgehen müssen, dann können wir uns von Folgerungen aus solchen Annahmen nicht mehr wünschen, als daß diese nicht noch ungewisser sein mögen. Eine Folgerung, die diese Eigenschaft hat, wollen wir eine *Wahrscheinlichkeitsfolgerung* (kurz: p-Folgerung, $\models_{\mathrm{p}}$) nennen:

$$A \models_{\mathrm{p}} B \text{ gdw } \forall \mathrm{p} : \overline{\mathrm{p}}(A) \geq \overline{\mathrm{p}}(B).$$

Die Verallgemeinerung zu Fällen mit mehr als einer Prämisse liegt auf der Hand: Die Ungewißheit der Konklusion darf nicht größer als die der Prämisse*n* sein. Aber wie messen wir die Wahrscheinlichkeit einer Pluralität $A_1, ..., A_n$ von Prämissen?

Für die Relation $\models$ logischer Folgerung gilt (unter elementaren Annahmen)

$$(1) \qquad A, B \models C \text{ gdw } A \wedge B \models C.$$

Das liegt daran, daß die Aussagen der Prämissenmenge $\{A, B\}$ genau dann wahr sind, wenn deren Konjunktion $A \wedge B$ wahr ist. Aber selbst wenn A und B im selben Grade ungewiß sind, ist dies im allgemeinen nicht der Grad der Ungewißheit von $A \wedge B$. (Ob die Münze in der geschlossen Hand oben Kopf oder Zahl zeigt, ist gleich ungewiß; daß beides oben liegt ist "maximal ungewiß", d.h. mit Sicherheit nicht der Fall.) Die Äquivalenz (1) wird also falsch, wenn wir logische ($\models$) durch Wahrscheinlichkeitsfolgerung ($\models_{\mathrm{p}}$) ersetzen. Wir wissen jedoch auch (vgl. den Beweis der nächsten Beobachtung), daß $\overline{\mathrm{p}}(A \wedge B)$ nie größer als $\overline{\mathrm{p}}(A) + \overline{\mathrm{p}}(B)$ sein kann, d.h.

$$(2) \qquad \overline{\mathrm{p}}(A) + \overline{\mathrm{p}}(B) \geq \overline{\mathrm{p}}(A \wedge B).$$

Aus PL und (1) folgt ferner ($\forall \overline{\mathrm{p}}$),

$$(3) \qquad \text{wenn } A, B \models C, \text{ dann } \overline{\mathrm{p}}(A \wedge B) \geq \overline{\mathrm{p}}(C),$$

und also nach (2)

$$(4) \qquad \text{wenn } A, B \models C, \text{ dann } \overline{\mathrm{p}}(A) + \overline{\mathrm{p}}(B) \geq \overline{\mathrm{p}}(C).$$

So haben wir die richtige Verallgemeinerung logischer Folgerung für den Fall nur wahrscheinlicher Prämissen und Konklusionen gefunden:

Def. $\models_{\mathrm{p}}$ $\qquad A_1, ..., A_n \models_{\mathrm{p}} B \text{ gdw } \forall \mathrm{p} : \overline{\mathrm{p}}(A_1) + \cdots + \overline{\mathrm{p}}(A_n)) \geq \overline{\mathrm{p}}(B).$

Wahrscheinlichkeitsschlüsse sind solche, in denen die Ungewißheit der Konklusion nie die summierte Ungewißheit der Prämissen übersteigen kann, gleichgültig wie ungewiß Prämissen und Konklusion sind. Jeder logisch gültige Schluß hat diese Eigenschaft, wie die nächste Beobachtung, (4) verallgemeinernd, zeigt.

Satz 9. (Adams' "Erster Hauptsatz" [3, p. 38].)
1. *Wenn $A_1, ..., A_n \models B$, dann $A_1, ..., A_n \models_{\mathrm{p}} B$.*
2. *Wenn $A_1, ..., A_n$ und B wahrheitswertig sind, dann gilt auch: Wenn $A_1, ..., A_n \models_{\mathrm{p}} B$, dann $A_1, ..., A_n \models B$.*

Beweis. *Ad 1.* Angenommen, $A_1, ..., A_n \models B$. Dann $\neg B \models \neg A_1 \vee \cdots \vee \neg A_n$ und also nach PL,

$$\mathrm{p}(\neg B) \leq \mathrm{p}(\neg A_1 \vee \cdots \vee \neg A_n).$$

Daraus folgt nach (P2′), daß

$$\mathrm{p}(\neg B) \leq \mathrm{p}(\neg A_1) + \cdots + \mathrm{p}(\neg A_n),$$

Also nach Definition $\overline{\mathrm{p}} = 1 - \mathrm{p}$,

$$\overline{\mathrm{p}}(B) \leq \overline{\mathrm{p}}(A_1) + \cdots + \overline{\mathrm{p}}(A_n).$$

Ad 2. Folgt aus dem Umstand, daß jede Wahrheitswertverteilung über geeignete Aussagen auch eine Wahrscheinlichkeitsverteilung ist. ■

Die Lotterie-Paradoxie. Man beachte, daß Argumente p-gültig sein können, auch wenn sie von jeweils einzeln hochwahrscheinlichen Prämissen auf eine ganz unwahrscheinliche Konklusion führen. Hier ist ein berühmtes Beispiel [168]: In einer Lotterie gebe es 1000 Lose. $L_1, \ldots, L_{1000}$ seien die Aussagen, daß das Los mit der Nummer $1 \ldots 1000$ gewinnt, und L stehe für $L_1 \vee \cdots \vee L_{1000}$, also für die Aussage, daß ein Los gewinnt. Ob ein bestimmtes Los L_i gewinnt, ist sehr ungewiß (vernünftigerweise $1 - 0,001$), jede Aussage $\neg L_i$ ist daher hochwahrscheinlich (0,999) – praktisch eine Gewißheit. (Wenn eine Aussage mit einer Wahrscheinlichkeit von 0,999 noch keine praktische Gewißheit sein sollte, dann erhöhe man im Beispiel die Anzahl der Lose.) Das Beispiel verführt zu einem paradox anmutenden Gedankengang. Wenn jede Annahme L_i praktisch sicher ist, dann sollte logisches Folgern aus der Gesamtheit dieser Annahmen nicht zu unsicheren Konklusionen führen. Aber aus $\neg L_1, \ldots, \neg L_{1000}$ folgt logisch die völlig unwahrscheinliche, tatsächlich falsche Konklusion $\neg L$, daß *kein* Los gewinnt. Wie kann das sein? Der Satz liefert eine Erklärung. Auch wenn einzelne Prämissen nur eine kleine Restungewißheit haben, so steigt die Ungewißheit einer Menge solcher Prämissen mit der Anzahl ihrer Elemente. Ist der Schluß gültig, so setzt die Ungewißheit der Konklusion eine obere Schranke für die Ungewißheit der Prämissenmenge (so der Satz). Wenn – wie in der Lotterie-Paradoxie – genügend viele Prämissen mit jeweils kleiner Ungewißheit zusammenkommen, so kann diese obere Schranke erreicht werden. Dann ist die Unsicherheit der Prämissen, als Gesamtheit aufgefaßt, gleich der Unsicherheit der Konklusion. Der Schluß ist in diesem Fall zwar p-gültig, aber es wäre nicht richtig ihn auszuführen. Täten wir es, dann würden wir aus Prämissen schließen, die, zusammengenommen betrachtet, maximal unsicher sind. (Genauso wie der Schluß von $\neg L_1, ..., \neg L_{1000}$ auf $\neg L$ zwar logisch gültig, aber nicht richtig ist, da die Prämissen nicht allesamt wahr sein können.)

Die Relation der p-Folgerung erfüllt die wichtigsten Bedingungen, die wir uns von Folgerungsrelationen erwarten; oder, etwas anders ausgedrückt, die Operation

$$C_{\mathrm{p}}(X) = \{A : X \models_{\mathrm{p}} A\}$$

erfüllt die charakteristischen Bedingungen einer Tarski'schen Konsequenzoperation:

Reflexivität	$X \subseteq C_{\mathrm{p}}(X);$
Monotonie	$X \subseteq Y \Rightarrow C_{\mathrm{p}}(X) \subseteq C_{\mathrm{p}}(Y);$
Transitivität	$C_{\mathrm{p}}(C_{\mathrm{p}}(X)) \subseteq C_{\mathrm{p}}(X).$

Was die Eigenschaft der Kompaktheit betrifft (wenn $X \subseteq C_{\mathrm{p}}(Y)$, dann gibt es eine endliche Teilmenge $Y' \subseteq Y$ mit $X \subseteq C_{\mathrm{p}}(Y')$), so ist diese trivial erfüllt. Da wir in endlichen Wahrscheinlichkeitsräumen arbeiten, so ist die Addition von Wahrscheinlichkeiten nur für eine endliche Anzahl von Summanden definiert. Also kann $\models_{\mathrm{p}}$ nur für endliche Prämissenmengen definiert sein.

Es wird nun Zeit, indikativische Konditionale wieder ins Bild zu bringen. Solche Konditionale können nicht Terme der logischen Folgerungsrelation

$$A_1, \ldots, A_n \models B \text{ gdw } \forall I : \min(I(A_1), \ldots, I(A_n) \leq I(B))$$

sein, denn sie sind keine zulässigen Argumente für Wahrheitswertverteilungen I. Aber da sie wahrscheinlichkeitsbewertet werden können, so können wir danach fragen, wie sie in der p-Folgerungsrelation,

$$A_1, \ldots, A_n \models_{\mathrm{p}} B \text{ gdw } \forall \overline{\mathrm{p}} : \overline{\mathrm{p}}(B) \leq \overline{\mathrm{p}}(A_1) + \cdots + \overline{\mathrm{p}}(A_n)$$

zueinander stehen.

Wir betrachten jetzt eine Sprache, in der die wahrheitsfunktionalen Zusammensetzungen um einen zweistelligen Junktor $\succ$ ergänzt sind. Von einem solchen Junktor erwarten wir normalerweise, daß er zur Bildung von Formeln so beiträgt, daß Formeln wie

$$(1)\ (A \succ B) \land C \quad \text{und} \quad (2)\ (A \succ B) \succ C$$

entstehen. Dabei müssen wir jedoch bedenken, daß die Formeln der Sprache als Terme der Relation $\models_{\mathrm{p}}$ verwendet werden sollen. Nach Def. $\models_{\mathrm{p}}$ steht die Relation zwischen Wahrscheinlichkeitsausdrücken. Die Formationsregeln einer Sprache mit $\succ$ sollten also so sein, daß nur legitime Argumente für

Wahrscheinlichkeitsfunktionen p als Formeln erzeugt werden können. Sind (1) und (2) solche Argumente?

Wir erinnern uns: Als Argumente einer Funktion p beschreiben Formeln Klassen von konkreten Ereignissen, Ereignistypen, wie wir gesagt haben. (In den klassischen Darstellungen der Wahrscheinlichkeitstheorie sind das die Ergebnisse, unter die Zufallsexperimente subsumiert werden können.) So beschreibt $A \wedge B$ die Klasse der Ereignisse, welche sowohl unter die Beschreibung A als auch unter die Beschreibung B fallen. Aber $A \succ B$ beschreibt gar keine Ereignisklasse, sondern allenfalls (so die Idee) das Verhältnis zweier solcher Klassen. Ein Ausdruck wie $\mathrm{p}(A \succ B)$ ist daher sinnlos, solange wir nicht festlegen, wie wir ihn verstehen wollen. Genau das tut AT: $\mathrm{p}(A \succ B)$ ist eine andere Art, sich auf den Wert $\mathrm{p}_A(B)$ zu beziehen. Aber dieser Weg, den Ausdruck $\mathrm{p}(A \succ B)$ zu eliminieren, ist nur gangbar, wenn das Konditional als Argument von p allein steht – auf $\mathrm{p}((A \succ B) \wedge C)$ ist AT nicht anwendbar. Die Formel (1) beschreibt also nichts und daher ist p darauf nicht anwendbar.

Wie steht es mit dem iterierten Konditional (2)? Sicher beschreibt auch diese Formel kein Ereignis; aber können wir den fragwürdigen Ausdruck $\mathrm{p}((A \succ B) \succ C)$ nicht mit Hilfe von AT zugunsten von $\mathrm{p}_{A \succ B}(C)$ eliminieren? Nach Def. p_A hätten wir dann aber

$$\mathrm{p}_{A \succ B} C = \frac{\mathrm{p}((A \succ B) \wedge C)}{\mathrm{p}(A \succ B)}$$

und somit im Zähler das gleiche Problem wie im Falle von (1).

Funktioniert die Elimination nicht zumindest im Falle von links iterierten Konditionalen wie

$$(3)\ A \succ (B \succ C),$$

in denen A, B, und C nicht selbst Konditionale enthalten? Das sieht zunächst besser aus, denn in diesem Fall ist

$$(*) \qquad \mathrm{p}(A \succ (B \succ C)) = \mathrm{p}_A(B \succ C) = (\mathrm{p}_A)_B(C).$$

Wir erinnern uns jedoch an dieser Stelle, daß die zweite dieser Gleichungen direkt verantwortlich ist für das Trivialitätsresultat von Lewis; vgl. $(*)$ auf p. 264. Wenn wir dem Resultat entgehen wollen, dürfen wir es zu dieser Gleichung nicht kommen lassen. Es gibt ein weiteres Argument gegen die Gleichung, welches unabhängig ist von der Trivialitätsdrohung. Dazu stellen wir zunächst fest, daß

$$(\mathrm{p}_A)_B(C) = \mathrm{p}_{A \wedge B}(C) = \mathrm{p}(A \wedge B \succ C).$$

Es folgt, daß unter der Relation der p-Folgerung die Äquivalenz

$$A \succ (B \succ C) \equiv_{\mathrm{p}} A \wedge B \in C$$

gelten muß. Die Folgerungsrichtung von rechts nach links nennt man *Exportation* (eines Konjunkts in eine Antezedensposition). Nun wünschen wir uns für $\succ$ unter p-Folgerung Modus Ponens:

$$A \succ B, A \models_{\mathrm{p}} B.$$

(Man erinnere sich daran, daß Konditionale die Funktion von Folgerungsfahrscheinen haben sollen!) Aber dann können wir so argumentieren:

(1)	$\models A \wedge B \to A$	Logik
(2)	$\models_{\mathrm{p}} A \wedge B \succ A$	(1), Beob. 9.1
(3)	$\models_{\mathrm{p}} A \succ (B \succ A)$	(2), Exportation
(4)	$A \succ (B \succ A), A \models_{\mathrm{p}} B \succ A$	Modus Ponens
(5)	$A \models_{\mathrm{p}} B \succ A$	(3),(4)

In (5) haben wir eine Sequenz hergeleitet, von der wir schon mehrfach festgestellt haben, daß sie keine gültige Folgerung für indikativische Konditionale ist. Die zu meidende Konklusion ist letzten Endes auf $(*)$ zurückführbar.

Diese Überlegungen schließen nicht aus, daß wir besondere Vereinbarungen für Formeln wie (1–3) als Argumente von p treffen, wie wir das ja auch schon durch AT für alleinstehende Konditionale getan haben. Aber ohne solche Vereinbarungen sind wir gut beraten, die Regeln für Formelbildung so zu gestalten, daß sie ohne weiteres legitime Argumente für Wahrscheinlichkeitsfunktionen produzieren.

In dieser Absicht beginnen wir mit dem Abschluß einer Menge von Atomen unter einer funktional vollständigen Menge Boole'scher Junktoren. Diese Menge nennen wir die *faktischen Formeln* der aus den Atomen generierten Sprache. Die Menge $\mathrm{FML}^{\succ}$ der *Formeln* ist dann so definiert:

Jede faktische Formel ist in $\mathrm{FML}^{\succ}$;
wenn A und B faktisch sind, dann ist $A \succ B$ in $\mathrm{FML}^{\succ}$;
$\mathrm{FML}^{\succ}$ sei die kleinste Menge, welche diese Bedingungen erfüllt.

Nach dieser Definition enthält die Sprache keine Formeln mit irgendwie eingebetteten Konditionalen. Soweit ein Konditional als Term der $\models_{\mathrm{p}}$-Relation vorkommen kann, ist es "flach", d.h. mit faktischem Antezedens und Konsequens.

Nun ist die Relation der p-Folgerung für Sprachen mit indikativischen Konditionalen definiert. Welche Folgerungen sind gültig und welche sind es nicht? Im folgenden stellen wir drei Verfahren vor, Antworten auf diese Frage zu finden.

Ersatzwahrheitstafeln. Konditionale haben keine Wahrheitswerte. Aber wir können so tun, als ob sie welche hätten. Eine einfache Weise, einige ungültige Instanzen der Relation $\models_{\mathrm{p}}$ zu identifizieren, besteht darin ihre Terme in wahrheitswertige Formeln umzuwandeln und dann auf logische Folgerung zu prüfen. Dazu definieren wir das materiale Gegenstück A^m einer Formel A so:

1. Wenn A faktisch ist, dann ist $A^m = A$;
2. wenn A ein Konditional $B \succ C$ ist, dann ist $A^m = B \to C$;
3. wenn $X = \{A_1, ..., A_n\}$, dann $X^m = \{A_1^m, ..., A_n^m\}$.

Beobachtung 10. *Wenn $X \models_{\mathrm{p}} A$, dann $X^m \models A^m$.*

Beweis. Unmittelbar aus Satz 9. ■

Das jetzt zu beschreibende Verfahren bedient sich ebenfalls der Simulation von Konditionalen durch ihre materialen Gegenstücke und erweist sich als richtiger und vollständiger Test auf Gültigkeit.

Es seien $A_1, ..., A_n, B$ Sätze der oben beschriebenen Sprache.

1. Konstruiere eine *Ersatzwahrheitstafel* für $A_1, ..., A_n, B$. Die Zeilen der Tafel wollen wir Bewertungen (v) nennen. Für faktische Formeln sind die Bewertungen klassisch, d.h. $v(\neg A) = 1$ gdw $v(A) = 0$ und $v(A \wedge B) = 1$ gdw $v(A) = v(B) = 1$. Für $A \succ B$ gilt

$$v(A \succ B) = \begin{cases} v(A \to B), \text{ wenn } v(A) \neq 0; \\ \text{undefiniert anderenfalls.} \end{cases}$$

2. Eine Bewertung v *bestätigt* eine Menge X von Sätzen genau dann, wenn (a) $\exists A \in X : v(A) = 1$ und (b) $\neg\exists A \in X : v(A) = 0$.

Satz 11. [3, p. 171] *$X \models_{\mathrm{p}} B$ genau dann, wenn $\exists Y \subseteq X$:*
(a) $Y^m \models B^m$, und
(b) $\forall v$: wenn v alle Sätze in Y bestätigt, dann $v(B) = 1$.

Beispiel. Wir wollen die folgenden Sequenzen prüfen:

$$B \models_{\mathrm{p}} A \succ B \tag{1}$$
$$A, B \models_{\mathrm{p}} A \succ B \tag{2}$$
$$A \succ B \models_{\mathrm{p}} A \to B \tag{3}$$
$$A \to B \models_{\mathrm{p}} A \succ B \tag{4}$$
$$A \succ B \models_{\mathrm{p}} \neg B \succ \neg A \tag{5}$$

Hier ist die Ersatzwahrheitstafel, die dem Zweck dienen wird:

A	B	$A \to B$	$A \succ B$	$\neg B \succ \neg A$
1	1	1	1	–
1	0	0	0	0
0	1	1	–	–
0	0	1	–	1

Alle Sequenzen bestehen offensichtlich den Test (a) des Satzes. Wir müssen also nur (b) untersuchen.

Ad (1). v_1 und v_3 bestätigen B, aber $v_3(A \succ B) \neq 1$. Nicht gültig.
Ad (2). Nur v_1 bestätigt $\{A, B\}$ und $v_1(A \succ B) = 1$. Gültig.
Ad (3). v_1 bestätigt $A \succ B$ und $v_1(A \to B) = 1$. Gültig.
Ad (4). Drei Bewertungen bestätigen $A \to B$, aber in den letzten beiden Zeilen hat $A \succ B$ nicht den Wert 1. Nicht gültig.
Ad (5). Wie in (3) bestätigt v_1 $A \succ B$, jedoch $v_1(\neg B \succ \neg A) \neq 1$. Nicht gültig.

Größengeordnete Euler-Diagramme. Das gerade beschriebene Verfahren ist sehr einfach und effizient, gibt aber keine Information darüber, wie die Wahrscheinlichkeitsverteilungen aussehen können, die dazu führen, daß eine Sequenz ungültig ist. Das ist aber eine wesentliche Information, um konkrete Gegenbeispiele zu Sequenzen zu konstruieren und auf ihre Plausibilität zu prüfen. Adams beschreibt daher ein weiteres Verfahren, das für p-Folgerung richtig und vollständig ist. Eine präzise Beschreibung dieses Verfahrens ist deutlich komplizierter als im Falle der Ersatzwahrheitstafeln. Wir wollen uns hier mit einer Skizze begnügen.

Euler-Diagramme stellen rein qualitative Verhältnisse zwischen Mengen dar.[37] Da wir Aussagen als Punktmengen auffassen können (Mengen von Wahrheitswertverteilungen, Zustandsbeschreibungen, möglicher Welten, etc.) finden Euler-Diagramme auch ihre offensichtliche Verwendung in der Aussagenlogik. Die Diagramme 1 und 2 stellen beide dasselbe dar: A und $\neg A$ schließen einander aus und erschöpfen zusammen den Bereich aller Möglichkeiten. Das Paar von Diagrammen macht deutlich, daß in Euler-Diagrammen eine bildhafte Information nicht genutzt wird: die *Größe* der Flächen. In Diagramm 1 ist A nicht “wahrer” als in Diagr. 2.

[37] Das tun auch Venn-Diagramme. Venn-Diagramme sind Euler-Diagramme, die eine zusätzliche Bedingung erfüllen. Adams’ Diagramme erfüllen diese Bedingung oft nicht. Adams verwendet also Euler-Diagramme, die in einigen Fällen zugleich Venn-Diagramme sind.

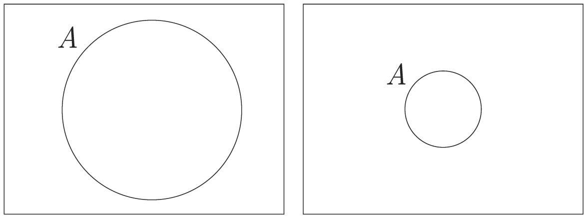


Diagramme 1 und 2

Da die Relation $\models_{\text{p}}$ auf einem Vergleich nicht zwischen Wahrheitswerten, sondern zwischen Wahrscheinlichkeitswerten beruht, bietet es sich an, die Größeninformation in Diagrammen zu nutzen, um Wahrscheinlichkeitsgrössen darzustellen. Wenn wir uns so die Kreisinhalte als proportional zu Wahrscheinlichkeiten vorstellen, dann geben die Diagramme 1 und 2 sehr verschiedene Situation wieder: Im einen Diagramm ist A wahrscheinlicher als $\neg A$, im anderen ist es umgekehrt. Es ist klar, daß in jedem Diagramm alle nicht überlappenden Flächeninhalte sich zum Flächeninhalt des Rechtecks summieren müssen. Jede Aufteilung des Rechtecks stellt also eine Wahrscheinlichkeitsverteilung dar.

In den beiden Diagrammen oben zeigt der A-Kreis den Anteil von A am gesamten Bereich an. ($|A|$ bezeichne den Teil eines Rechtecks, der die Aussage A darstellt.) Eine konditionale Wahrscheinlichkeit $\text{p}_A(B)$ gibt die Wahrscheinlichkeit von B an unter der Annahme, daß wir uns im A-Bereich befinden. Im Diagramm sehen wir uns dazu zunächst den Teil von $|B|$, der in $|A|$ liegt, d.h. $|A \wedge B|$ an. Dieser Teil kann beispielsweise wie in Diagr. 3 oder wie in Diagr. 4 ausfallen.

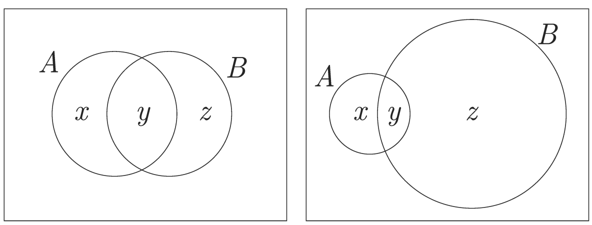


Diagramme 3 und 4

Dann fragen wir, welchen Anteil $y = |A \wedge B|$ an $x + y = |A|$ hat, d.h. wir dividieren y durch $x + y$ (wie es die Definition konditionaler Wahrscheinlichkeit vorsieht). Je größer y im Verhältnis zu x ist, umso höher ist $\text{p}_A(B)$; Diagr. 3 zeigt offenbar einen kleineren Wert von $\text{p}_A(B)$ an als Diagr. 4. Anders als die Wahrscheinlichkeit einer faktischen Formel wird die eines Konditionals also nicht durch eine abgezirkelte Fläche im Rechteck dargestellt. Es ist vielmehr ein Verhältnis solcher Flächen, welches die Wahrscheinlichkeit eines Konditionals festlegt. Das ist ein bildhafter Ausdruck der These, daß Konditionale nichts beschreiben und daher ihre Wahrscheinlichkeit nicht die Wahrscheinlichkeit sein kann, daß eine Beschreibung zutrifft.

Aus der diagrammatischen Methode läßt sich eine Definition von Folgerung entwickeln, die mit der Relation der p-Folgerung koinzidiert; Details sind in [2] und [3] dargestellt. Wir betrachten die Methode hier nur, um Gegenbeispiele zu Folgerungsbehauptungen zu finden und als Heuristik, um für die Gültigkeit von Folgerungen zu argumentieren. Beides wollen wir anhand einiger Sequenzen illustrieren.

Für die Widerlegungen einiger Folgerungsbehauptungen, wird die Betrachtung von Diagr. 4 genügen. Hier ist (wie auch in Diagr. 3) $|A| = x+y$, $|B| = y + z$, und $|A \wedge B| = y$. Danach ist

$$\mathrm{p}_A(B) = \frac{y}{x+y} \quad \text{und} \quad \mathrm{p}_B(A) = \frac{y}{y+z}.$$

Nun ist in Diagramm 4 der Anteil von y an $y + z$, der $\mathrm{p}_A(B)$ wiedergibt, deutlich kleiner als $|B| = y + z$. Damit gibt das Diagramm ein Rezept ab für Gegenbeispiele zur Behauptung

$$A \models_{\mathrm{p}} B \succ A. \tag{1}$$

Ein solches Gegenbeispiel ist die Verteilung mit

$$x = 0,2,\ y = 0,1,\ \text{und } z = 0,5.$$

Dann ist

$$\mathrm{p}(B) = y + z = 0,6 \quad \text{und} \quad \mathrm{p}_A(B) = \frac{y}{y+x} = \frac{0,1}{0,3} = 0,\overline{3}.$$

Das abstrakte Gegenbeispiel läßt sich leicht in ein konkretes ummünzen, indem wir plausible Einsetzungen für A und B finden. Beispiel:

A: Die Eintracht spielt in Unterzahl.
B: Die Eintracht gewinnt.

Aus einer guten Chance, daß die Eintracht gewinnt, folgt keine mindestens ebenso gute Chance, daß sie gewinnt, wenn sie in Unterzahl spielt.

Auch Kontraposition,

$$\neg A \succ B \models_{\mathrm{p}} \neg B \succ A, \tag{5}$$

wird durch Diagr. 4 und die Werteverteilung widerlegt. Wir rechnen:

$$\mathrm{p}_{\neg A}B = \frac{|\neg A \wedge B|}{|\neg A|} = \frac{z}{1-(x+y)} = \frac{0,5}{0,7} \approx 0,71;$$
$$\mathrm{p}_{\neg B}A = \frac{A \wedge \neg B}{\neg B} = \frac{x}{1-(y+z)} = \frac{0,2}{0,4} = 0,5.$$

(An diesem Beispiel sieht man, daß es mit dem heuristischen Wert der Diagramme nicht immer weit her ist. In vielen Fällen wird es einfacher sein, sich zuerst eine widerlegende Werteverteilung zu überlegen, um danach ein dieser Verteilung mehr oder weniger entsprechendes Diagramm zu konstruieren.) Einsetzungen für A und B illustrieren das in plausibler Weise. So ist beispielsweise die Aussage

$\neg A \succ B$: Wenn die Eintracht in voller Stärke spielt, dann gewinnt sie.

deutlich wahrscheinlicher als die vermeintliche Konklusion:

$\neg B \succ A$: Wenn die Eintracht nicht gewinnt, dann spielt sie in Unterzahl.

Diagramm 4 zeigt ebenfalls wie

(4) $$A \rightarrow B \models_{\mathrm{p}} A \succ B$$

falsch sein kann. Die Wahrscheinlichkeit von $A \rightarrow B$ ist mindestens so hoch wie die von B, welche, wie das Bild zeigt, höher ist als die von $A \succ B$. Mit den oben festgesetzten Werten ist $\mathrm{p}(A \succ B) = 0,\overline{3}$ und $\mathrm{p}(A \rightarrow B) = \mathrm{p}(\neg A \vee B) = (1-x+y)+y = 1-x = 0,8$. Als konkretes Gegenbeispiel zur Kontraposition mag wieder das Eintracht-Beispiel dienen. Man betrachte

$A \rightarrow B$: Die Eintracht spielt in voller Stärke oder gewinnt.
$A \succ B$: Wenn die Eintracht in Unterzahl spielt, dann gewinnt sie.

Das materiale Konditonal mag eine hohe Wahrscheinlichkeit haben, weil B (oder $\neg A$) für sehr wahrscheinlich gehalten wird. Jemand, der die Dinge so sieht, verpflichtet sich nicht dazu unter der Annahme der Unterzahl den Gewinn für mindestens ebenso wahrscheinlich zu halten.

Dagegen ist die umgekehrte Sequenz

(3) $$A \succ B \models_{\mathrm{p}} A \rightarrow B$$

(Modus Ponens für $\succ$) gültig. Wir betrachten das Diagramm 3, das wir nun aber so verstehen wollen, daß es für alle Wahrscheinlichkeitsverteilungen noch offen sein soll. (Daß die beiden Kreise gleich groß sind und sich überlappen, soll jetzt also nicht heißen, daß wir nur Verteilungen über A und B betrachten, unter denen die Aussagen gleich wahrscheinlich sind und ihre Konjunktion nicht gleich Null ist.) Die Größen der Kreise sind gewissermaßen nur provisorisch angezeigt und wir fragen jetzt danach, ob sie so festgelegt werden können, daß ein Gegenbeispiel zur Sequenz entsteht. Wie zuvor ist

$$\mathrm{p}(A \succ B) = \frac{y}{x+y} \quad \text{und} \quad \mathrm{p}(A \to B) = (1 - x + y) + y = 1 - x.$$

Es ist nun einfach zu zeigen, daß für alle x, y, z $(x + y > 0)$,

$$\frac{y}{x+y} \leq 1 - x.$$

Dazu erinnern wir uns, daß $\mathrm{p}_A(B) = 1 - \mathrm{p}(\neg B)$. Für die Werte in unserm Diagramm bedeutet das:

$$x/(x + y) = 1 - y/(x + y).$$

Da $x+y \leq 1$, so ist $x \leq x/(x+y)$. Also $x \leq 1-y/(x+y)$ und somit $y/(x+y) \leq 1 - x$ wie gewünscht. Es kann also keine Festlegung der Kreisgrößen geben so, daß ein Gegenbeispiel zur Modus Ponens-Sequenz angezeigt wird.

Schließlich wollen wir noch eine Sequenz in drei Variablen diagrammatisch widerlegen:

$$A \succ B, B \succ C \models_{\mathrm{p}} A \succ C. \tag{6}$$

Die folgenden zwei Diagramme zeigen zwei verschiedene Weisen, wie indikativische Konditionale nicht transitiv verkettet sein können.

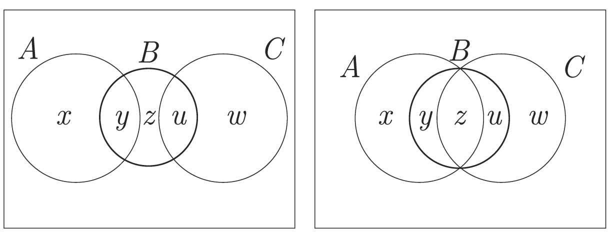


DIAGRAMME 5 und 6.

Diagr. 5 zeigt einen positiven Anteil von B an A (d.h. von y an $x + y$) und von C an B (d.h. von u an $y + z + u$); beide Prämissen haben also eine positive Wahrscheinlichkeit. Aber C hat gar keinen Anteil an A, d.h. $\mathrm{p}_A(C) = 0$. Eine Wahrscheinlichkeitsverteilung, die dieser Beschreibung entspricht, ist schnell gefunden; z.B. $x = w = 0,2$ und $y = z = u = 0,1$. Für eine konkrete Instantiierung von A, B und C benötigen wir $\mathrm{p}(A \succ B) \neq 0$, $\mathrm{p}(B \succ C) \neq 0$ und $A \wedge C \leftrightarrow \perp$. Hier ist ein Beispiel: [38]

A: Die Eintracht schießt genau drei Tore.
B: Die Eintracht schießt höchstens drei Tore.
C: Die Eintracht schießt weniger als drei Tore.

$A \succ B$ ist zwar keine logische Wahrheit (denn indikativische Konditionale beschreiben auch keine logischen Tatsachen), aber genauso wahrscheinlich wie eine logische Wahrheit. Die meisten Spiele der Eintracht enden mit höchstens zwei Toren. Wenn ich das weiß und daher C eine hohe Wahrscheinlichkeit zubillige, dann gibt die Annahme B keinen Anlaß, von diesem Wert abzuweichen. Also hat $B \succ C$ ebenfalls eine hohe Wahrscheinlichkeit. Aber $A \succ C$ ist völlig unwahrscheinlich.

Das aus Diagr. 5 abgeleitete Gegenbeispiel ist etwas ungewöhnlich, da es von einer maximal wahrscheinlichen Prämisse und einer maximal unwahrscheinlichen Konklusion ausgeht. Eine bei weitem gewöhnlichere Situation, in der Transitivität fehlgeht, zeigt Diagr. 6. Hier hat B einen großen Anteil an A, nämlich $y+z$; C hat einen ebenso großen Anteil an B, nämlich $z+u$; jedoch hat C nur einen relativ kleinen Anteil an A, nämlich z ($< y+z$ und $< z+u$). Das klassische Beispiel, auf das diese Beschreibung zutrifft ist dieses:

A: Es gehen Lawinen ab.
B: Der Boden ist mit Schnee bedeckt.
C. Ich fahre Ski.

Es ist klar, was den Schluß von $A \succ B$ und $B \succ C$ auf $A \succ C$ vereitelt. Während $B \succ C$ für sich betrachtet hochwahrscheinlich sein mag, so würden wir nicht auf C schließen wollen, wenn das Antezedens B durch Ablösung von A gewonnen wurde; anders gesagt: $A \wedge B \succ C$ hat eine dramatisch geringere Wahrscheinlichkeit als $B \succ C$. Zum einen erklärt das, warum Gegenbeispiele zu Transitivität immer auch solche gegen Prämissenverstärkung,

$$B \succ C \models_{\mathrm{p}} A \wedge B \succ C, \tag{7}$$

sind – wie der Leser anhand der Beispiele nachprüfen sollte. [39] Zum anderen deutet es darauf hin, daß die folgende Sequenz (Vorsichtige Transitivität)

[38] Nach [3, p. 127]. Vgl. auch [148, p, 48]: “Wenn ich gewinne, bin ich Kandidat”, “Wenn ich kandidiere, verliere ich”.

[39] Es ist klar, daß Transitivität mit Prämissenverstärkung fallen muß. Denn da $A \wedge B \succ A$ gültig ist, wäre der Schluß von $A \succ C$ auf $A \wedge B \succ C$ zulässig, wenn $\succ$ transitiv wäre.

gültig ist:

(8) $$A \succ B, A \wedge B \succ C \models_{\mathrm{p}} A \succ C.$$

Die Adams-Logik AL. Neben den zwei bisher vorgestellten Beschreibungen der Menge gültiger p-Folgerungen, gibt Adams [3, Kap. 7] noch eine rein syntaktische Beschreibung der p-gültigen Folgerungen an. Wir präsentieren dieses System hier etwas anders als in [4], nämlich als Sequenzenkalkül. (Die Buchstabenetiketten stammen von Adams, die Kurzbeschreibungen sind der neueren Literatur angeglichen.)

Das System **AL** von Adams

A, B, C seien faktische Formeln, $\vdash_0$ sei Ableitbarkeit in der klassischen Logik. X, Y seien endliche Mengen von Formeln und das Komma links von $\triangleright$ deute Vereinigung an (so, daß $X, Y = Y, X$ und $X, X = X$).

Verdünnung $$\frac{X \triangleright A}{X, Y \triangleright A}$$ Schnitt $$\frac{X \triangleright A \quad \forall B \in X : Y \triangleright B}{Y \triangleright A}$$

G. Reflexivität $$A \triangleright A$$

LC. Logische Konsequenz $$\frac{A \vdash_0 B}{X \triangleright A \succ B}$$

CF. $\top$-Ein/-Aus $$\frac{X \triangleright A}{X \triangleright \top \succ A} \qquad \frac{X \triangleright \top \succ A}{X \triangleright A}$$

EA. Äquiv. im Antezed. $$\frac{A \equiv_0 B \quad X \triangleright A \succ C}{X \triangleright B \succ C}$$

RT. Kumulative Trans. $$\frac{X \triangleright A \succ B \quad Y \triangleright A \wedge B \succ C}{X, Y \triangleright A \succ C}$$

DA. Disj. im Antezed. $$\frac{X \triangleright A \succ C \quad Y \triangleright B \succ C}{X, Y \triangleright A \vee B \succ C}$$

AR. Vorsichtige Monot. $$\frac{X \triangleright A \succ B \quad Y \triangleright A \succ C}{X, Y \triangleright A \wedge B \succ C}$$

Der Beweis des folgenden Satzes ist in [4, §7.6] skizziert.

SATZ 12. *Die ableitbaren Sequenzen sind genau die* P*-gültigen Folgerungen:* $\models_{\mathrm{P}} \; = \; \triangleright$.

Es folgen einige beispielhafte Herleitungen von Sequenzen und Regeln. Zunächst die "starke Zentrierungs-Sequenz" $A, B \triangleright A \succ B$:

(1)	$A \triangleright A$	G
(2)	$B \triangleright B$	G
(3)	$A \triangleright \top \succ A$	1, CF
(4)	$B \triangleright \top \succ B$	2, CF
(5)	$A, B \triangleright A \wedge \top \succ B$	3,4, AR
(6)	$A, B \triangleright A \succ B$	5, EA

Die Modus Ponens-Regel $\dfrac{X \triangleright A \quad Y \triangleright A \succ B}{X, Y \triangleright B}$:

(1)	$X \triangleright A$	
(2)	$Y \triangleright A \succ B$	
(3)	$X \triangleright \top \succ A$	1, CF
(4)	$Y \triangleright A \wedge \top \succ B$	2, EA
(5)	$X, Y \triangleright \top \succ B$	3,4, RT
(6)	$X, Y \triangleright B$	5, CF

Wenn wir in der Ableitung A für X und $A \succ B$ für Y einsetzen, dann generiert diese die Modus Ponens-Sequenz $A, A \succ B \triangleright B$. Ebenfalls eine Art Modus Ponens drückt die Sequenz $A \succ B \triangleright A \rightarrow B$ aus:

(1)	$A \succ B \vdash A \succ B$	G
(2)	$\triangleright A \wedge B \succ (A \rightarrow B)$	LC: $A \wedge B \vdash_0 A \rightarrow B$
(3)	$A \succ B \triangleright A \succ (A \rightarrow B)$	1,2, RT
(4)	$\triangleright \neg A \succ (A \rightarrow B)$	LC: $\neg A \vdash_0 A \rightarrow B$
(5)	$A \succ B \triangleright A \vee \neg A \succ (A \rightarrow B)$	3,4, DA
(6)	$A \succ B \triangleright A \rightarrow B$	5, EA: $\top \equiv_0 A \vee \neg A$, CF

Konditionale sind transitiv *via* $\vdash_0$, $\dfrac{X \triangleright A \succ B \quad B \vdash_0 C}{X \triangleright A \succ C}$:

(1)	$X \triangleright A \succ B$	
(2)	$B \vdash_0 C$	
(3)	$A \wedge B \vdash_0 C$	2
(4)	$\triangleright A \wedge B \succ C$	3, LC
(5)	$X \vdash A \succ C$	1,3, RT

Schließlich noch die zu DA "duale" Regel $\dfrac{X \triangleright A \succ B \quad Y \triangleright A \succ C}{X, Y \triangleright A \succ B \wedge C}$:

(1)	$X \triangleright A \succ B$	
(2)	$Y \triangleright A \succ C$	
(3)	$\triangleright B \wedge C \succ B \wedge C$	LC
(4)	$X \triangleright A \wedge C \succ B \wedge C$	1,3, RT
(5)	$X, Y \triangleright A \wedge A \succ B \wedge C$	2,4, RT
(6)	$X, Y \triangleright A \succ B \wedge C$	5, EA

Die entsprechende Sequenz $A \succ B, A \succ C \triangleright A \succ A \wedge B$ wird mit $X = A \succ B$ und $Y = A \succ C$ genauso hergeleitet.

All diese logischen Prinzipien sind uns, wenn auch in etwas anderer Form, aus der Logik der kontrafaktischen Konditionale bekannt. So entsprechen beispielsweise die Sequenzen $A \succ B \triangleright A \to B$ und $A \wedge B \triangleright A \succ B$ in offensichtlicher Weise den Schemata der schwachen bzw. starken Zentrierung, (CW) $A \sqsupset B \to (A \to B)$ und (CS) $A \wedge B \to (A \sqsupset B)$. Die starke Zentrierung, CS, ist das charakteristische Schema für die Logik (VC) der Sphärensysteme. Für Schemata wie (K) $(A \sqsupset (B \to C)) \to ((A \sqsupset B) \to (A \sqsupset C))$ dagegen können wir keine entsprechende Sequenz erwarten, denn $\sqsupset$-Formeln simulierende $\succ$-Formeln können ja nicht Teil komplexerer Formeln sein. Mit dieser Einschränkung stellen wir so eine weitgehende Übereinstimmung in den logischen Prinzipien für kontrafaktische und indikativische Konditionale fest. Tatsächlich ist die Übereinstimmung vollkommen, wie der nächste Satz zeigt.

Es sei $\models_S$ die Relation gültiger Folgerung in Sphärenmodellen (oder den entsprechenden variabel relationalen Modellen).

> $X \models_S A$ gdw in beliebigen Sphärenmodellen gilt: An jedem Punkt, der alle Formeln in X wahr macht, ist auch A wahr.

Es sei $\mathrm{FML}^{\sqsupset}$ die Menge der Formeln einer Sprache mit kontrafaktischen Konditionalen; $\mathrm{FML}^{\succ}$ ist, wie auf p. 272 definiert, die Formelmenge einer Sprache mit indikativischen Konditionalen. Die Formeln in $\mathrm{FML}^{\sqsupset}$ können verschachtelte Konditionale enthalten; die in $\mathrm{FML}^{\succ}$ sind allesamt “flach”, d.h. rein faktisch oder von der Form $A \succ B$ mit A und B faktisch. Um hier auf einen gemeinsamen Nenner zu kommen, betrachten wir das flache Fragment von $\mathrm{FML}^{\sqsupset}$, bestehend aus rein wahrheitsfunktionalen Formeln oder solchen der Form $A \sqsupset B$ mit wahrheitsfunktionalen Teilformeln A und B. Das flache Fragment von $\mathrm{FML}^{\sqsupset}$ unterscheidet sich von $\mathrm{FML}^{\succ}$ nur noch in der Bezeichnung des Konditionals – hier $\succ$, dort $\sqsupset$. Diesen typographischen Unterschied wollen wir vernachlässigen und einfach von “flachen Formeln” sprechen.

SATZ 13. (Stalnaker [272].) *Es sei X eine Menge flacher Formeln und A eine flache Formel. Dann $X \models_S A$ gdw $X \models_{\mathrm{P}} A$, und $X \vdash_{\mathbf{VC}} A$ gdw $X \triangleright A$.*

10. Rückblick: Die Logik der Konditionale

Der letzte Satz hält ein erstaunliches Resultat fest. Haben wir nicht gesagt, daß sich kontrafaktische und indikativische Konditionale semantisch grundlegend unterscheiden? Die einen sind mögliche Welten-Konditionale, die anderen sind es nicht. Folgerung wurde für die einen über Sphären (oder variabel relationale Modelle) definiert, für die anderen über Wahrscheinlichkeitsfunktionen. Die Modellierungen könnten kaum weiter auseinanderliegen – und doch koinzidieren sie am Ende im Hinblick auf die als gültig ausgezeichneten Folgerungen (im flachen Fragment der Sprache).

Warum sind die logischen Prinzipien für kontrafaktische und indikativische Konditionale einander so ähnlich, ja stimmen vielleicht sogar völlig überein? Sicher gibt es Unterschiede zwischen Wahrscheinlichkeitsschlüssen einerseits und Schlüssen andererseits, die sich aus der Betrachtung einander ähnelnder Welten ergeben. Aber offenbar gibt es eine Stufe der Abstraktion, bei der diese Unterschiede verschwinden und nur noch Gemeinsames im Blick bleibt. Einen ersten Hinweis auf eine enge strukturelle Ähnlichkeit geben die Bedingungen für Wahrheit bzw. hohe Wahrscheinlichkeit der Konditionale.

$\sqsupset$: An einem Punkt a ist $A \sqsupset B$ wahr, wenn B wahr ist in einem Bereich minimal von a abweichender Punkte, an denen die Annahme A gilt:
$I(A \sqsupset B, a) = 1$ gdw $R_{[\![A]\!]}(a) \subseteq [\![B]\!]$.

$\succ$: Unter einer Funktion P ist $A \succ B$ hochwahrscheinlich (h), wenn B hochwahrscheinlich ist unter einer minimal von P abweichenden Funktion, die unter der Annahme A steht:
$\mathrm{P}(A \succ B) \approx h$ gdw $\mathrm{P}_A(B) \approx h$.

Das gemeinsame Schlüsselelement ist offenbar die *minimale Abweichung* von einem Ausgangspunkt so, daß das Antezedens des Konditionals wahr bzw. den Maximalwert 1 hat. Zwar sind die Ausgangspunkte sehr verschieden – hier Wahrheitswertverteilungen, dort Wahrscheinlichkeitsverteilungen –, aber der Effekt ist schließlich derselbe: Dadurch, daß wir nicht jede, sondern nur bevorzugte Alternativen betrachten, werden bestimmte Annahmen konstant gehalten, die nun gemeinsam mit dem Antezedens – im Erfolgsfall – das Konsequens erzwingen (d.h. seine Wahrheit bzw. hohe Wahrscheinlichkeit). Ohne diese Zusatzannahmen wäre das Antezedens nicht hinreichend für das Konsequens. Dabei kann eine Verstärkung des Antezedens dazu führen, daß die für das Konsequens nötigen Zusatzannahmen außer

Kraft gesetzt werden. Das erklärt die Nichtmonotonie im Antezedens, d.h. die Ungültigkeit des Schlusses von $A \sqsupset B$ auf $A \wedge C \rightarrow B$ bzw. von $A \succ B$ auf $A \wedge C \succ B$, sowie weiterer Schlußfiguren.

Soweit die allgemeine Erklärung, warum kontrafaktische und indikativische Konditionale auf der Ebene ihrer logischen Eigenschaften einander sehr ähnlich sind. Obwohl die Erklärung einfach ist, stellen sich genauere Nachweise der engen Beziehung zwischen den beiden Arten von Konditionalen – d.h. Repräsentationsresultate für verschiedene Modellierungen – als recht aufwendig dar. Das gilt auch für den Beweis von Satz 13. Oft laufen solche Abbildungen über die Theorie der Theorienrevision, die wir in Kapitel VIII behandeln werden. Das liegt daran, daß die Operation der Revision (einer Theorie) im wesentlichen die Operation eines minimalen Eingriff ist. So thematisiert die Theorie der Theorienrevision das wichtigste Modellierungselement der Konditionallogik.[40] Das gilt ebenso – wie wir sehen werden – für eine weitere logische Theorie, die eng mit der Konditionallogik verwandt ist: die Theorie anfechtbaren (oder nicht-monotonen) Schließens, die im nächsten Kapitel behandelt wird.

[40] Hinter der Rede von einer “minimalen Abweichung” versteckt sich tatsächlich eine ganze Familie von Relationen; vgl dazu insbesondere [196].

V.

PARAKONSISTENTE LOGIK

In diesem und dem nächsten Kapitel geht es nicht um Dinge, welche die klassische Logik *nicht* behandelt, wie die intensionalen Operationen der Modallogik. Es geht vielmehr um die Möglichkeit, daß die klassische Logik einiges von dem, was sie behandelt, *falsch* behandelt. Der Verdacht fällt hier insbesondere auf die klassische Theorie der Negation, der Implikation und der logischen Konsequenz. Während wir in den vorhergehenden Kapiteln im wesentlichen mögliche Erweiterungen der klassischen Logik betrachtet haben, soll es nun also um Alternativen, d.h. um sogenannte nicht-klassische Logiken gehen. [1]

1. Inkonsistenz und Trivialität

In der Erzählung "Die Bibliothek von Babel" [35] stellt der argentinische Schriftsteller Jorge Luis Borges eine "Universalbibliothek" vor: Sie enthält alle Bücher mit 410 Seiten (40 Zeilen pro Seite, 80 Zeichen pro Zeile), die aus einem Alphabet von 25 Zeichen (inkl. Leerzeichen, Komma und Punkt)

[1] Die hier unterstellte Unterscheidung zwischen Erweiterung und Alternative ist alles andere als unproblematisch. Darauf hat inbesondere Quine wiederholt hingewiesen; siehe z.B. [237, Kap. 6]. Welche Art von Tatsachen könnte man anführen, um den Unterschied zwischen einer anderen Interpretation eines Junktors und der Interpretation eines anderen Junktors zu markieren? Wir werden später sehen, daß sich zumindest einige logische Systeme, die zunächst als Alternativen zur klassischen Logik auftreten, auch als Erweiterungen der klassischen Logik um neue Junktoren darstellen lassen.

erzeugt werden können.[2] Bücher, welche nur den Buchstaben A (gefolgt von Leerzeichen) enthalten oder die Buchstaben M C V immer wieder, und scheinbar sinnlos, kombinieren, stehen dort ebenso in den Regalen, wie ein Buch, welches mit der so unwahrscheinlichen Zeichenfolge beginnt, in der wir den Text von Borges' *Ficciones* erkennen können (mit den Druckfehlern der ersten Auflage von 1944).

Einen kleinen Teil der Bibliothek füllt das *Corpus Teutonicum*, d.i. die Gesamtheit aller möglichen Bücher, die dem Lexikon und der Grammatik der deutschen Sprachen folgen. Neben den vielen deutschen Übersetzungen der *Ficciones* gibt es im *Corpus Teutonicum* auch ein Buch mit dem Titel "Summa", daß aus nur zwei Sätzen besteht (gefolgt von Leerzeichen): "Schnee ist weiß. Schnee ist nicht weiß." Über die *Summa* behaupten klassische Logiker etwas erstaunliches, nämlich daß das Büchlein denselben Inhalt wie das gesamte *Corpus Teutonicum* hat. Aus logischer Sicht, so behauptet der Klassiker, ist die aus zwei Sätzen bestehende *Summa* dem mächtigen *Corpus Teutonicum* gleichwertig. Das begründet der Klassiker wie folgt.

Der Inhalt eines Buches ergibt sich aus dem, was darin geschrieben steht. Wenn in einem Buch der Satz steht

(58) Benjamín Otálora ist um 1891 neunzehn Jahre alt,

dann gehört es zum Inhalt des Buches, daß Benjamín Otálora 1891 noch am Leben ist und auch daß jemand um 1891 neunzehn Jahre alt ist. Diese beiden Sachverhalte folgen logisch aus dem Satz (58). Jeder, der diesen einen Satz des Buches liest und versteht, kann diese Sachverhalte dem Satz ohne weiteres entnehmen. Der Inhalt eines Buches besteht also aus dem, was darin wörtlich steht, zusammen mit dem, was daraus logisch geschlossen werden kann, d.h. der Information, die dem Text zu entnehmen ist. Der Inhalt *ist* die Information, die das Buch enthält.

Nun ist der Satz (58) auf den Seiten der *Summa* zwar nicht zu finden. Dennoch gehört er aus Sicht der klassischen Logik zum Inhalt des Buches – genau wie jeder andere Satz des *Corpus Teutonicum*. Denn aus einem beliebigen Widerspruch, d.h. zwei Sätzen der Form A und $\neg A$, folgt jeder

[2] Borges phantasiert präzise: Die Bibliothek besteht aus sechseckigen Sälen, in denen jeweils 960 Bücher eingestellt sind; in 30 Regalen jeweils 32 Bücher. Da jedes Buch eine endliche Permutation über einer endlichen Basis (25) ist, so ist die Anzahl möglicher Bücher endlich. Jedes Buch enthält $n = 410 \times 40 \times 80 = 1.312.000$ Zeichen, so daß es 25^n mögliche Bücher gibt. Im Internet kursieren geschätzte Vergleiche zwischen der Größe der Bibliothek von Babel und der Größe des sichtbaren Universums – wobei sich dieses als recht klein im Vergleich mit jenem erweist. Borges Erzählung – 1941 erstmals veröffentlicht – spinnt ein Gedankenexperiment weiter, das schon Kurd Laßwitz [169] in der Erzählung "Die Universalbibliothek" (1904) beschrieben hat.

beliebige Satz B, also insbesondere auch (58). Das ist eine elementare Eigenschaft der klassischen Folgerungsrelation, *ex falso quodlibet* genannt:[3]

(efq) Wenn A und $\neg A$, dann B.

Also gehört (1), wie jeder andere deutsche Satz, zum Inhalt der *Summa*. Umgekehrt ist die *Summa* natürlich Teil des *Corpus Teutonicum*. Also haben *Summa* und *Corpus Teutonicum* denselben Inhalt.

Irgendwie vermag dieser Gedankengang nicht so recht zu überzeugen: das riesige *Corpus Teutonicum* in nur zwei kurzen Sätzen zusammengefaßt? Es will sich hier nicht der Eindruck einstellen, daß die klassische Logik uns zu einer Entdeckung über das Verhältnis der *Summa* zum *Corpus Teutonicum* verholfen hat. Stattdessen drängt sich die Frage auf: Was läuft falsch in der klassischen Logik, wenn sie zur Behauptung gezwungen ist, daß ein Paar von widersprüchlichen Sätzen äquivalent zur Menge aller Sätze ist? Das Argument beruhte wesentlich auf zwei Prämissen:

1. Abschluß: Was aus einer Menge von Sätzen logisch folgt, gehört zu dessen Inhalt.
2. *Ex falso quodlibet*: Aus einer widersprüchlichen Menge von Sätzen folgt logisch jeder beliebige Satz.

Die logischen Theorien, die wir in diesem Kapitel betrachten werden, richten ihre kritische Aufmerksamkeit auf die zweite Prämisse. (Klassiker werden die Aufmerksamkeit auf die erste Prämisse lenken wollen, um dem ungläubigen Staunen die Spitze zu nehmen.) Das anti-klassische Argument ist sehr einfach: Erstens, so behauptet der Nicht-Klassiker, wüßten wir sehr gut zu unterscheiden zwischen Widersprüchlichkeit einerseits und Trivialität, d.h. der unterschiedslosen Behauptung beliebiger Sätze, andererseits. Nun gibt es viele Unterschiede, die uns klar und deutlich sind, die aber logisch bedeutungslos und deshalb in logischer Hinsicht keine sind. Deshalb argumentiert der nicht-klassische Logiker, zweitens, daß eine Theorie, die den *logischen* Unterschied zwischen Inkonsistenz und Trivialität aufhebt, unbefriedigend sei.

[3] Da diese Eigenschaft *elementar* ist, können wir auch gleich einen möglichen Einwand gegen die gerade gegebene Bestimmung des Inhalts eines Buches beantworten. Enthält ein Zettel mit den Dedekind-Peano-Axiomen der Arithmetik die gesamte Zahlentheorie, die sich daraus entwickeln läßt? In einem bestimmten Sinne von “enthalten” bzw. “Inhalt” würden wir das nicht sagen wollen. Ein Buch über Zahlentheorie, das sich an Mathematiker wendet, enthält – in einem vertrauten Sinn von “enthalten sein” – mehr “Stoff” als ein Schulbuch. Aber was ein Leser aufgrund *elementarer* Schlüsse dem Wortlaut eines Textes entnehmen kann, würden wir auch in einem eingeschränkteren Sinne zum Inhalt des Textes zählen. Der Schluß nach *ex falso quodlibet* ist, klassisch gesehen, elementar.

Die erste Behauptung ist wenig strittig. Kontroversen entzünden sich an der zweiten Behauptung. Hier muß der Nicht-Klassiker zeigen, daß es interessante inkonsistente Theorien gibt. Solche Theorien würden natürlich jedes Interesse verlieren, sobald man sie unter klassischer Folgerung abschließt und so trivialisiert. Den Inhalt einer Theorie, die uns interessiert, lernen wir nur kennen, indem wir ihn logisch erschließen. Das gilt auch für inkonsistente Theorien. Wenn wir den nicht-trivialen Inhalt einer solchen Theorie kennenlernen wollen, dann muß das Schließen einer logischen Theorie folgen, die nicht die klassische sein kann. Die Frage ist also: Was ist eine interessante inkonsistente Theorie und gibt es dafür Beispiele?

Eine inkonsistente Theorie wäre nicht interessant, wenn wir sogleich wüßten, wie sie zu reparieren sei – einmal vorausgesetzt, daß Inkonsistenz einen "Defekt" anzeigt. Interessante inkonsistente Theorien sind also derart, daß es keine eindeutig bessere, konsistente Version der Theorie gibt. Solange das so ist, darf die inkonsistente Version einer Theorie unser Interesse beanspruchen. Abhängig vom Menü vorliegender Alternativen mag die inkonsistente Theorie sogar die beste Option sein.

Klassiker könnten an dieser Stelle einwerfen, daß Inkonsistenz ein so großer Schaden an einer Theorie sei, daß diese im Vergleich zu jeder ernsthaften konsistenten Alternative unterliegen *müsse*. Natürlich darf dieser Einwurf sich nicht auf die Annahme stützen, daß inkonsistente Theorien schlecht sind, weil sie trivial sind. Genau diese Annahme bestreitet der Nicht-Klassiker ja. Aber der Klassiker könnte darauf hinweisen, daß ein Widerspruch niemals wahr sein kann und deshalb jede widersprüchliche Theorie den Makel sicherer Falschheit trägt. Die konsistenten Alternativen mögen verschroben oder ontologisch inflationär sein; aber im Gegensatz zur inkonsistenten Theorie haben sie den unbestreitbaren Vorzug, möglicherweise wahr zu sein.

Der Nicht-Klassiker kann auf zwei, sehr unterschiedliche Weisen antworten. Er kann, erstens, bestreiten, daß Widersprüche nicht wahr sein können. Das würde inkonsistente Theorien sogleich vom Makel der sicheren Falschheit befreien. Die Behauptung, daß es wahre Sätze der Form $A \wedge \neg A$ gibt, ist die charakteristische These des *Dialethismus.* Die dialethische These scheint mindestens genauso gewagt wie die klassische Gleichsetzung von Inkonsistenz und Trivialität. Insofern wäre es wünschenswert, wenn der Nicht-Klassiker eine zweite, weniger verwegene Antwort zur Hand hätte. Hier ist eine *pragmatische* Antwort: Der Widerspruch in einer Theorie zeigt nur an, daß ein *Teil* der Theorie falsch ist. Es ist nun gerade die Aufgabe einer nicht-klassischen Logik, diesen inkonsistenten Teil inferenziell zu isolieren und zum zuverlässigen Teil der Theorie einen zuverlässigen logischen Zugang zu schaffen. Es gibt *ab initio* keinen Grund zu der Annahme, daß eine mit Vorsicht gehandhabte inkonsistente Theorie schlechter sein muß als ihre konsistenten Alternativen. Vielmehr kann es durchaus so sein, daß die inkonsistente Theorie einfacher und erklärungsstärker ist als konsistente Konkurrenten, und daß wir im Umgang mit der Theorie sicher feststellen können, in welchen Ableitungen wesentlich von konsistenten bzw. inkonsistenten Annahmen Gebrauch gemacht wird.

Es seien hier beispielhaft drei Theorien skizziert, die interessant und inkonsistent sind. In allen drei Fällen ist die Frage, welche konsistente Theorie über den jeweiligen Gegenstand die beste sei, in hohem Grade strittig. Ferner ist die jeweilige inkonsistente Theorie – soweit nicht konsequent unter klassischer Folgerung abgeschlossen – schon aufgrund ihrer Einfachheit eine ernsthafte Option. Im Falle der ersten zwei Beispiele (Mengen und Wahrheit betreffend) deutet sogar alles darauf hin, daß wir mit genau diesen inkonsistenten Theorien täglich und meist arglos arbeiten. Die drei Beispiele sind darüberhinaus auch gute Kandidaten, die dialethische These zu belegen.[4]

Mengenlehre. Für jede Bedingung A scheint es eine Menge y zu geben, die genau die Elemente enthält, welche die Bedingung erfüllen. Das ist das sogenannte "naive" Komprehensionsschema der Mengenlehre:

$$(*) \qquad \exists y \forall x (x \in y \leftrightarrow Ax).$$

Das Schema $(*)$ versichert uns unter anderem, daß eine Menge r existiert, deren Elemente die Bedingung $\bullet \notin r$ erfüllt (die Russell-Menge). Daher gilt:

$$\forall x (x \in r \leftrightarrow x \notin r).$$

Wenn wir nun r für x einsetzen, dann erhalten wir

$$(\dagger) \qquad r \in r \leftrightarrow r \notin r.$$

Entweder ist nun $r \in r$ oder es ist nicht so, d.h. $r \notin r$. In beiden Fällen folgt unmittelbar der Widerspruch aus $(\dagger)$: $r \in r$ und $r \notin r$.

Jede Mengenlehre, die das völlig natürlich scheinende Schema $(*)$ enthält, ist also inkonsistent. Solange Mengenlehre nicht selbst Gegenstand ist, setzen Mathematiker immer so etwas wie das naive Komprehensionsschema

[4] Weitere Beispiele inkonsistenter Theorien und vielleicht inkonsistenter Gegenstandsbereiche finden sich insbesondere in den Arbeiten von Graham Priest [222, 224, 225, 227].

(∗) voraus und arbeiten damit im Rahmen einer inkonsistenten Theorie.[5] Konsistente Einschränkungen von (∗) sind mühsam, erscheinen *ad hoc*, und ihre deduktive Kraft in manchen Fällen nicht ausreichend oder ungeklärt. Eine Alternative könnte darin bestehen, mit (∗) weiter zu arbeiten und durch eine Einschränkung der Logik das Abgleiten in eine triviale Theorie zu verhindern.

Wahrheitstheorie. Über den Begriff der Wahrheit wird in der Philosophie viel gestritten. Aber dieses eine ist sicher: Ein Satz ist wahr, wenn es so ist, wie er es sagt. "Schnee ist weiß" ist genau dann wahr, wenn Schnee weiß ist. Auch diese Binsenweisheit können wir in einem Schema ausdrücken. Sei S ein beliebiger deutscher Satz und $\ulcorner S \urcorner$ ein Name für diesen Satz. Dann gilt:

$$\ulcorner S \urcorner \text{ ist wahr gdw } S. \tag{T}$$

Nun sei L der Satz "$\ulcorner L \urcorner$ ist nicht wahr". Wenn wir L für S in (T) einsetzen, dann erhalten wir

$$\ulcorner L \urcorner \text{ ist wahr gdw } L.$$

Aber $L = (\ulcorner L \urcorner$ ist nicht wahr). So erhalten wir

$$\ulcorner L \urcorner \text{ ist wahr gdw } \ulcorner L \urcorner \text{ nicht wahr ist.}$$

Wie im Falle der naiven Mengentheorie folgt daraus, daß $\ulcorner L \urcorner$ zugleich wahr und nicht wahr ist. Das Resultat können wir so zusammenfassen: Jede Theorie der Wahrheit, welche von dem Schema (T) ausgehend die Wahrheitsbedingung für einen Satz wie L bestimmen will, ist widersprüchlich. Wenn wir dem Widerspruch entgehen wollen, dann müssen wir entweder das sehr plausible Schema (T) aufgeben oder Sätze wie L aus dem Skopus der Wahrheitstheorie einer Sprache ausschließen. Im ersten Fall verlieren wir eine wesentliche Einsicht in den Begriff der Wahrheit (die sogenannte Disquotierungseigenschaft); im zweiten Fall wird unsere Wahrheitstheorie unvollständig. Wenn wir beides nicht wollen, dann könnte es das Beste sein, den Widerspruch zu akzeptieren und eine Logik zu wählen, die unsere Wahrheitstheorie dennoch nicht trivialisiert.

---

[5] Selbst in manchen frühen Darstellungen der Mengenlehre ist nicht der Hauch einer Beunruhigung spürbar. So schreibt zum Beispiel Felix Hausdorff im letzten Satz des Vorworts zur zweiten Auflage (1927) seiner *Mengenlehre*: "Zu einer Diskussion über Antinomien und Grundlagenkritik habe ich mich jetzt ebensowenig wie damals [1914] entschließen können." Auch in der Auflage letzter Hand, aus dem Jahre 1935, änderte sich daran nichts.

Vagheit. Ob eine Ansammlung von Sandkörnern ein Haufen ist, mag manchmal nicht ganz klar sein. Aber drei Dinge wissen wir sicher über Haufen:

1. Ein Sandkorn ist kein Haufen.
2. Sandhaufen entstehen nicht "plötzlich", indem man zu einem Nichthaufen ein einziges Sandkorn hinzulegt (Toleranzprinzip für Haufen).
3. Eine Million Sandkörner (richtig zusammengelegt) sind ein Haufen.

Beginnen wir mit der ersten Gewißheit: Ein Sandkorn ist kein Haufen. Erste Anwendungen des Toleranzprinzips führen zu den nicht minder gewissen Feststellungen, daß zwei und drei Sandkörner auch keine Haufen sind. Ganz allgemein gilt nach dem Toleranzprinzip, daß wenn n Sandkörner kein Haufen sind, dann sind auch $n+1$ Sandkörner kein Haufen. Wenn wir das Toleranzprinzip immer wieder anwenden, gelangen wir zum Schluß:

4. Eine Million Sandkörner sind kein Haufen.

Die Aussagen 3 und 4 widersprechen einander. Was sollen wir daraus schließen? Wenn wir einmal die erste und die dritte Gewißheit als Grunddaten über die Bedeutung von "Haufen" ausklammern, dann könnten wir versuchen den Schluß abzuwenden, indem wir entweder das Toleranzprinzip oder die Verkettung von Modus Ponens-Schlüssen – das einzige logische Prinzip, das hier zur Anwendung kommt – kritisch unter die Lupe nehmen. Das sind ersichtlich schwierige Wege, da das Toleranzprinzip ein Charakteristikum vager Begriffe zu sein scheint und Modus Ponens wohl zum Bedeutungskern von "wenn-dann" gehört. Wir könnten aber auch dem logischen Schließen aus plausiblen Prämissen zur Konklusion 4 mutig folgen. Vage Begriffe erwiesen sich so als inkonsistent. Dann gibt es zwei Optionen. Der Klassiker würde sagen: Inkonsistente Begriffe sind notwendigerweise leer; es gibt also gar keine Haufen.[6] Der Dialethiker würde sagen: Natürlich gibt es Haufen; das sind Gegenstände, die unter einen inkonsistenten Begriff fallen. Das Paar der Sätze 3 und 4 wären demnach ein Beispiel eines wahren Widerspruchs, aus dem wir geichwohl nicht auf Beliebiges schließen wollen.[7]

Dies sind Beispiele von Theorien, die aufgrund einer ihrem jeweiligen Gegenstand inhärenten Schwierigkeit widersprüchliche Thesen enthalten. Aus dialethischer Perspektive sind diese Theorien möglicherweise wahr. Neben solchen Schaufensterstücken des Dialethismus gibt es viele weniger spektakuläre und sehr alltägliche inkonsistente Theorien. Vermutlich gilt für die meisten von uns, daß die Gesamtheit unserer Überzeugungen inkonsistent ist. Die Gründe sind vielfältig. Wir speichern neue Informationen ab,

[6] Diese Sicht vertreten Wheeler [286] und Unger [282].
[7] Hyde [145] entwickelt eine Theorie vager Prädikate aus parakonsistenter Sicht.

ohne alte, widersprechende zu löschen. Wir rufen in verschiedenen Kontexten verschiedene Überzeugungen ab – möglicherweise aus verschiedenen Quellen – und geraten nie in Kontexte, in denen sich die Unvereinbarkeit dieser Überzeugungen zeigen könnte. Manchmal ist es sehr schwierig, die Inkonsistenz von Überzeugungen zu erweisen. (Frege war jahrelang davon überzeugt, daß seine Version der Mengenlehre konsistent sei.) Schließlich tolerieren wir manchmal leichte Inkonsistenzen in unseren Überzeugungen – jedenfalls solange uns keine naheliegende Reparatur einfällt und wir meinen, daß die Inkonsistenzen praktisch folgenlos bleiben. In allen diesen Fällen würden wir sagen: Diese Überzeugungssysteme sind zwar inkonsistent; aber sie haben dennoch logisch verschiedenen Inhalt und sind insbesondere nicht gleichermaßen trivial. Aus der Sicht der klassischen Logik wäre das jedoch falsch: Jedes inkonsistente Überzeugungssystem hat denselben logischen Gehalt, nämlich den trivialen Gehalt eines Überzeugungssystem in dem schlechthin alles geglaubt wird.

2. Parakonsistenz

Eine *Theorie* (im logischen Sinne) ist eine Menge von Sätzen, die unter logischer Konsequenz abgeschlossen ist. Sei $\mathrm{Cn}(X)$ die Menge der logischen Konsequenzen einer Satzmenge X. Dann ist X genau dann eine Theorie, wenn

$$X = \mathrm{Cn}(X).$$

Eine Theorie X ist *inkonsistent*, wenn es zwei Sätze A und $\neg A$ gibt, mit $A \in \mathrm{Cn}(X)$ und $\neg A \in \mathrm{Cn}(X)$; X ist *trivial*, wenn $\mathrm{Cn}(X)$ die Menge aller Sätze der zugrundeliegenden Sprache ist. Ein Theorie ist *parakonsistent*, wenn sie inkonsistent und nicht trivial ist.

Gegeben eine Logik **L**, so bestimmen wir die Menge $\mathrm{Cn}(X)$ der logischen Konsequenzen (im Sinne von **L**) über die Ableitbarkeitsrelation $\vdash$ von **L** so:

$$A \in \mathrm{Cn}(X) \text{ gdw } X \vdash A \text{ in } \mathbf{L}.$$

Wenn Cn und $\vdash$ in dieser Beziehung zueinander stehen, dann ist Cn eine *Konsequenzoperation über* **L**. (Siehe dazu auch den Abschnitt 2.4 im Einleitungskapitel.) Eine Logik nennen wir *parakonsistent*, wenn sie eine Konsequenzoperation Cn definiert, die parakonsistente Theorien zuläßt – es also Satzmengen X gibt, so daß $\mathrm{Cn}(X)$ parakonsistent ist.

Wie eine parakonsistente Logik aussehen kann, erfahren wir am besten, indem wir uns vor Augen führen, warum in der klassischen Logik parakonsistente Theorien nicht möglich sind. Hier ist eine klassische Ableitung eines beliebigen Satzes B aus widersprüchlichen Annahmen A und $\neg A$: [8]

Erste Ableitung.

$$\dfrac{\dfrac{A}{A \vee B}\,\vee\text{-Einf.} \qquad \neg A}{B}\,\text{DS}$$

Alternativ können wir die Ableitung auch im Lemmon-Stil notieren: [9]

1	(1)	A	Annahme
1	(2)	$A \vee B$	1 $\vee$-Einf.
3	(3)	$\neg A$	Annahme
1,3	(4)	B	2,3 DS

Die Abkürzung "DS" steht für die Schlußfigur des *disjunktiven Syllogismus* im Sinne der folgenden Regel:

$$\frac{\neg A \quad A \vee B}{B}.$$

Die Ableitung zeigt – am deutlichsten im Lemmon-Stil –, daß wir den DS hier im starken Sinne einer Ableitungsregel verwendet haben: [10]

$$\frac{X \vdash \neg A \quad Y \vdash A \vee B}{X, Y \vdash B}$$

Gegeben die Reflexivität von $\vdash$, folgt aus der Regel DS unmittelbar die Sequenz

DS. $$\neg A, A \vee B \vdash B.$$

Und gegeben die Schnittregel, gilt auch die Umkehrung. Unter allgemein geltenden Annahmen über $\vdash$ sind Regel- und Sequenzenform also äquivalent.

Im Prinzip kann die erste Ableitung an jedem Schritt aufgehalten werden. Nur im Vorübergehen sei erwähnt, daß in den sogenannten *nicht-additiven Logiken* (siehe [276]) der Schritt (2), d.h. die $\vee$-Einführung, zurückgewiesen

[8] Das ist das sogenannte Lewis-Argument, benannt nach dem amerikanischen Logiker C.I. Lewis.

[9] Siehe die Erläuterungen in der Einleitung.

[10] D.h. die Regel DS ist eine abgeleitete, nicht bloß zulässige Regel. Siehe dazu Kap. I.2.

wird. Aber das ist eine Option, die nur unter sehr speziellen und kontroversen Annahmen über logisches Folgern begründet werden kann. Die meisten parakonsistenten Logiken erheben Einspruch gegen die Gültigkeit des disjunktiven Syllogismus im Sinne der Regel DS bzw. – äquivalent – im Sinne der Sequenz sDS. Wir werden gleich sehen, warum das ein guter Ansatzpunkt ist.

Dagegen spricht zunächst, daß der disjunktive Syllogismus in einem engen Zusammenhang mit Modus Ponens steht. Die Prämissen des Modus Ponens, A und $A \to B$, sind zumindest klassisch äquivalent zu A und $\neg A \vee B$, denn es gilt ja klassisch die Äquivalenz

$$(A \to B) \leftrightarrow (\neg A \vee B).$$

Parakonsistente Logik steht daher vor der Wahl, entweder diese Äquivalenz aufzugeben (genauer: die Richtung von rechts nach links) oder mit DS auch Modus Ponens zurückzuweisen.

Eine zweite klassische Ableitung von B aus A und $\neg A$ verweist auf die Möglichkeit eines zunächst überraschend anmutenden Ansatzes zu einer parakonsistenten Logik.

Zweite Ableitung.

$$\cfrac{\cfrac{A \quad \neg A}{A \wedge \neg A}\,\text{Adjunktion} \qquad A \wedge \neg A \to B}{B}\,\text{MP}$$

Oder so:

1	(1)	A	Annahme
2	(2)	$\neg A$	Annahme
1,2	(3)	$A \wedge \neg A$	1,2 Adjunktion
	(4)	$A \wedge \neg A \to B$	EFQ
1,2	(5)	B	3,4 MP

Diese Ableitung könnte man nicht erst im Schritt (4), sondern schon im Schritt (3) anhalten, indem man die Regel der Adjunktion (auch “$\wedge$-Einführung” genannt),

ADJ. $$\frac{X \vdash A \quad Y \vdash B}{X, Y \vdash A \wedge B},$$

bzw. die folgende äquivalente Sequenz ablehnt:

sADJ. $$A, B \vdash A \wedge B.$$

Eine Logik, in der die Adjunktion nicht gültig ist, würde auch die Instanz

$$\frac{A,\ \neg A}{A \wedge \neg A}$$

ablehnen und damit den Schluß von A und $\neg A$ auf B im zweiten Argument aufhalten können; sie wäre, im Sinne der Definition, parakonsistent.

Soweit sind das nur technische Maßnahmen, einige klassische Ableitungen scheitern zu lassen. Daß es solche Interventionsmöglichkeiten gibt, ist nicht *per se* interessant. Wie gut sind diese Interventionen motiviert? Letztlich möchten wir, daß sich die Ungültigkeit des disjunktiven Syllogismus oder der Adjunktion auf natürliche Weise daraus ergibt, daß wir die Möglichkeit nicht-trivialer inkonsistenter Theorien ernstnehmen und plausible Interpretationen dieser Möglichkeit anbieten.

3. Ein nicht-adjunktiver Ansatz: Diskussive Logik

Der Einspruch gegen die zweite Ableitung, der sich gegen das Prinzip der Adjunktion richtet, fußt auf einer sehr bestimmten Vorstellung davon, wie Widersprüche entstehen können. Wir betrachten hier Adjunktion zunächst in der umgangssprachlichen Formulierung:

> Aus A und B folgt $A \wedge B$.

Wenn das Zeichen $\wedge$ die Konjunktion ("und") repräsentieren soll, dann ist dieses Prinzip zwingend. Aber das folgende, ähnliche Prinzip ist sicher nicht empfehlenswert:

> Wenn jemand A behauptet und jemand B behauptet, dann behauptet jemand $A \wedge B$.

Ulla sagt: "Das Brentano-Bad öffnet morgen um sieben"; Peter sagt, daß es erst um acht Uhr öffne. Welche Meinungen werden in diesem kurzen Gespräch vertreten? Daß das Brentano Bad morgen öffnet (darin stimmen beide überein), daß eine städtische Einrichtung um sieben Uhr öffnet (das folgt aus Ullas Beitrag). Dies und mehr können wir als Ergebnis des Gesprächs festhalten. ("Ergebnis" in einem schwachen, vorläufigen Sinne, der ungelöste Kontroversen einschließt.) Aber zu den angebotenen Meinungen gehört sicher nicht, daß eine Badeanstalt um sieben *und* um acht Uhr öffnet. Für die Summierung der Ansichten verschiedener Parteien ist Adjunktion also kein guter Schluß.

Zeuge einer kontroversen Diskussion zu sein, ist ein einfaches Modell, wie wir in vielen Fällen zu unseren Überzeugungen kommen. Aus einer Quelle erfahren wir A; aus einer anderen Quelle erfahren wir $\neg A$. In Kontexten, in denen die Erinnerung an jene Quelle naheliegt, schließen und handeln wir aufgrund der Überzeugung A; in anderen Kontexten kommt uns eher $\neg A$ in den Sinn. Aus keiner Quelle jedoch haben wir $A \wedge \neg A$ erfahren. Diese, konjunktive Information steht uns in dem gerade beschriebenen Sinne gar nicht zur Verfügung.

Der polnische Logiker Stanisław Jaśkowski hat versucht diese Situation in einer Reihe logischer Theorien – die er "diskussive" (manchmal auch "diskursive") Logik nannte – zu modellieren; siehe [155, 156], sowie [57, 284]. Dazu bediente er sich der Modallogik. Den Operator $\Diamond$ verstand er als einen Existenzquantor, wonach $\Diamond A$ ausdrücken sollte: A ist der Fall "nach Meinung [mindestens] eines der Diskussionsteilnehmer" oder auch "nach [mindestens] einer bestimmten Auffassung der in A vorkommenden Termini" [156, p. 149]. Am einfachsten stellen wir uns dazu die Klasse der (universalen) **S5**-Modelle vor und interpretieren die Punkte des Modells als Diskussionsteilnehmer.

> *Erinnerung* (vgl. Kap. III.9). In einem (universalen) **S5**-Modell mit Punktemenge W ist $\Diamond A$ an einem Punkte $a \in W$ genau dann *wahr*, wenn es einen Punkt $b \in W$ gibt, so daß A am Punkte b wahr ist – kurz:
>
> $$a \models \Diamond A \text{ gdw } \exists b : b \models A.$$
>
> Ein Modell macht eine Formel A (bzw. Formelmenge X) genau dann wahr, wenn A (bzw. jede Formel in X) an jedem Punkt des Modells wahr ist:
>
> $$\mathcal{M} \models A \text{ gdw } \forall a \text{ in } \mathcal{M}:\ a \models A.$$
>
> Eine **S5**-*gültige* Formel ist eine Formel, die in allen **S5**-Modellen wahr ist. Schließlich *folgt* eine Formel A aus einer Formelmenge X ($X \models A$), wenn jedes Modell, das alle Formeln in X wahr macht, auch die Formel A wahr macht; also
>
> $$X \models A \text{ gdw } \forall \mathcal{M}:\ \mathcal{M} \models X \Rightarrow \mathcal{M} \models A.$$
>
> Eine Folgerung ist **S5**-gültig, wenn der Quantifikationsbereich in der Äquivalenz die Menge der **S5**-Modelle ist.

Wir legen nun als Formelmenge $\mathrm{FML}^{\Diamond}$ die der modalen Sprachen $\mathcal{L}^{\Diamond}$ zugrunde, d.h. den Abschluß einer Menge von Atomen unter Boole'schen Kombinationen sowie dem Operator $\Diamond$. Gegeben die Definition einer **S5**-gültigen Formel, sei die Menge der in **D5** *gültigen Formeln* so definiert:

Def. $$\mathbf{D5} = \{A \in \mathrm{FML}^{\Diamond}:\ \models \Diamond A \text{ in } \mathbf{S5}\}.$$

Und gegeben die Definition **S5**-gültiger Folgerung, bietet sich diese Definition *gültiger Folgerung* in **D5** an:

Def. $$X \models_{\mathbf{D5}} A \text{ gdw } \Diamond X \models_{\mathbf{S5}} \Diamond A,$$

wobei $\Diamond X = \{\Diamond A : A \in X\}$.

> *Anmerkung zur Benennung.* Jaśkowski nennt **S5** "$\mathbf{M_2}$" und die darauf aufbauende diskussive Logik "$\mathbf{D_2}$". Die Tradierung des historischen Namens ist offensichtlich "unlogisch", gegeben die übliche Benennung des zugrundeliegenen modalen Systems. Was hier als **D5** definiert ist, nennt da Costa [57] übrigens **J**. Tatsächlich entspricht **D5** nicht ganz dem System $\mathbf{D_2}$ von Jaśkowski in [156]: **D5** ist eine Teilmenge von FML$^{\Diamond}$, während $\mathbf{D_2}$ als Teilmenge einer anderen Formelmenge definiert ist. Siehe dazu unten die Definition des Systems $\mathbf{D5}^d$ ($\approx \mathbf{D_2}$).

Bildlich gesprochen, sind Modelle mit ihren Punkten wie Diskussionen mit ihren Teilnehmern. In diesem Bild ist eine Aussage A gültig, wenn es keine Diskussion gibt, in der nicht jemand A vorbringt – A also ein unvermeidlicher Beitrag ist. Eine Aussage B folgt aus A, wenn jede Diskussion, in der A vertreten wird, auch eine ist, in der jemand B vertritt – A also unweigerlich B nach sich zieht.

Die Folgerungsrelation von **D5** ist parakonsistent im Sinne der Definition auf p. 292, denn

$$A, \neg A \models_{\mathbf{D5}} B$$

ist nicht gültig, da $\Diamond A, \Diamond\neg A \models \Diamond B$ in **S5** offensichtlich nicht gilt. Andererseits ist aber $\Diamond(A \wedge \neg A) \models \Diamond B$ gültig in **S5**, weshalb $A \wedge \neg A \models B$ in **D5** gültig ist!

Ebenso gilt auch Modus Ponens nicht in der Form

$$A, A \to B \models_{\mathbf{D5}} B,$$

wohl aber wenn die Annahmen konjungiert sind: $A \wedge (A \to B) \models_{\mathbf{D5}} B$. Die Logik **D5** unterscheidet also zwischen den Annahmenmengen $\{A, B\}$ und $\{A \wedge B\}$. Das stimmt durchaus mit der eingangs beschriebenen Idee überein, nach der eine diskussive Logik nicht adjunktiv sein sollte. Da der Schluß von der Prämissenmenge $\{\Diamond A, \Diamond B\}$ auf die Konklusion $\Diamond(A \wedge B)$ ungültig ist (sei $B = \neg A$ für eine beliebige kontingente Aussage A), so gilt in **D5**

nicht: $$A, B \models_{\mathbf{D5}} A \wedge B.$$

Aus einem Paar von Formeln kann nicht auf deren Konjunktion geschlossen werden. Tatsächlich, haben Mengen von Formeln – sofern sie nicht modal sind – überhaupt keinen synergetischen Effekt für den Zweck des Schließens, wie die nächste Beobachtung zeigt.

Beobachtung 1. *Wenn $X \cup \{B\}$ eine Menge nichtmodaler Formeln ist, dann $X \models_{\mathbf{D5}} B$, nur wenn es ein $A \in X$ gibt mit $A \models_{\mathbf{D5}} B$.*[11]

Beweis. Ohne wesentliche Einschränkung der Allgemeinheit betrachten wir den Fall $A, B \models_{\mathbf{D5}} C$. Wir kontraponieren und nehmen an, daß $A \not\models_{\mathbf{D5}} C$ und $B \not\models_{\mathbf{D5}} C$, d.h.

$$\Diamond A \not\models_{\mathbf{S5}} \Diamond C \quad \text{und} \quad \Diamond B \not\models_{\mathbf{S5}} \Diamond C.$$

(Zu zeigen: $\Diamond A, \Diamond B \not\models_{\mathbf{S5}} \Diamond C$.) Dann gibt es zwei Modelle $\mathcal{M}_A = (W_A, I_A)$ und $\mathcal{M}_B$ so, daß

(a) in $\mathcal{M}_A$: $(\exists x \in W_A)\ I_A(A, x) = 1$ und $(\forall x \in W_A)\ I_A(C, x) = 0$;
(b) in $\mathcal{M}_B$: $(\exists x \in W_B)\ I_B(B, x) = 1$ und $(\forall x \in W_B)\ I_B(C, x) = 0$.

Nun sei $\mathcal{M} = (W_A \cup \{b\}, I)$ so, daß (1) $I = I_A$ für FML $\times\ W_A$, und (2) $I(B, b) = 1$ während $I(C, b) = 0$. Nach (b) gibt es eine solche Interpretation von B und C an einem Punkt b. Ferner ist Bedingung (2) konsistent mit (1). Denn da A, B und C nicht modal sind, ist I_A unabhängig von $I(B, b)$ und $I(C, b)$, und umgekehrt. Dann gilt

(c) in $\mathcal{M}$: $\exists x \in W_A : I_A(A, x) = 1$,
$I(B, b) = 1$ und
$\forall x \in W_A \cup \{b\} : I(C, x) = 0$.

Also widerlegt $\mathcal{M}$ die Sequenz $\Diamond A, \Diamond B \models_{\mathbf{S5}} \Diamond C$. ∎

Obwohl die beobachtete Superkompaktheit eine merkwürdige Eigenschaft für eine Folgerungsrelation ist, so könnte man sie unter der Diskussionsinterpretation noch annehmbar finden. Denn eine Menge nichtmodaler Formeln auf der linken Seite der Relation $\models_{\mathbf{D5}}$ faßt ja nur zusammen, was in einer Diskussion vorgebracht wurde. Nach der Beobachtung folgt aus einer solchen Zusammenfassung nicht mehr, als was logisch den einzelnen Beiträgen entnommen werden kann. Ob das richtig ist, ist nicht ohne weiteres zu beurteilen; jedenfalls aber ist es nicht offensichtlich abwegig.

Alle klassischen Folgerungen aus einer Formel sind auch diskussiv gültig. Denn, angenommen B folgt aus A klassisch, dann gilt $\Diamond A \models \Diamond B$ in **S5** und also $A \models B$ in **D5**. Wie wir oben gesehen haben, gilt insbesondere $A \wedge \neg A \models B$ in **D5**. Unter der diskussiven Interpretation bedeutet das, daß wenn ein Beitrag zur Diskussion einen Widerspruch enthält, dann ist der Beitrag trivial. Das paßt nicht gut zur Idee einer parakonsistenten Logik. Die diskussive Logik **D5** gibt sich hier vielmehr als eine *klassische* logische Theorie einer Diskussionssituation zu erkennen. Tatsächlich ist der Theorembestand von **D5** einfach eine Erweiterung der klassischen Logik.

[11] Eine etwas stärkere Behauptung (die Beobachtung gälte auch, wenn B einstufig modal ist) findet sich in [229, p. 161]. Sie wird jedoch durch ein Gegenbeispiel in [284, p. 241] widerlegt.

BEOBACHTUNG 2. **KL** $\subseteq$ **S5** $\subseteq$ **D5**.

BEWEIS. **KL** $\subseteq$ **S5** ist offensichtlich. Wenn $A \in$ **S5**, dann auch $\Diamond A \in$ **S5** und also nach Definition $A \in$ **D5**. ∎

D5 kann sehr einfach wie in da Costa [57] axiomatisiert werden:

$$\Box A, \text{ für jedes Theorem } A \text{ von } \mathbf{S5}$$

$$\frac{\Box(A \to B) \quad \Box A}{\Box B}\,\text{R1}$$

$$\frac{\Box A}{A}\,\text{R2} \qquad \frac{\Diamond A}{A}\,\text{R3} \qquad \frac{\Box A}{\Box\Box A}\,\text{R4}$$

Sei $\vdash$ die Ableitungsrelation im obigen axiomatischen System. Dann gilt:

SATZ 3. $\vdash A$ *gdw* $A \in$ **D5**.

BEWEIS. (LR) Induktion über die Länge eines Beweises im obigen System. Wenn A ein Axiom ist, dann $A = \Box B$ mit $B \in$ **S5**. Es folgt, daß $\Box B \in$ **S5**. Da $\Box B \to \Diamond\Box B \in$ **S5**, so folgt ferner, daß $\Diamond\Box B \in$ **S5** und also $\Box B\,[= A] \in$ **D5**. — Für jede Regel $X \Rightarrow A$ genügt es zu zeigen, daß die Regel $\Diamond X \Rightarrow \Diamond A$ jeweils in **S5** zulässig ist. Die Regeln entsprechen einfachen **S5**-Theoremen:

R1: $\Diamond\Box(A \to B) \to (\Diamond\Box A \to \Diamond\Box B)$;
R2: $\Diamond\Box A \to \Diamond A$;
R3: $\Diamond\Diamond A \to \Diamond A$;
R4: $\Diamond\Box A \to \Diamond\Box\Box A$.

(RL) Angenommen $A \in$ **D5**. Dann gilt (nach Definition) $\Diamond A \in$ **S5** und also $\vdash \Box\Diamond A$. Anwendung von R2 ergibt $\vdash \Diamond A$, worauf wir R3 anwenden können mit der Konklusion $\vdash A$. ∎

Wie Beobachtung 1 zeigt, gibt es in **D5** keine echten Schlüsse auf nichtmodale Konklusionen aus mehr als einer nichtmodalen Prämisse. Das liegt nicht nur an der Ungültigkeit der Adjunktion, sondern auch daran, daß die Regel Modus Ponens, die Ableitungen vorantreibt, in **D5** für die gewöhnliche Implikation nicht gelten kann. Wenn die Beobachtung für modale Prämissenmengen und Konklusionen nicht gilt, dann sollte es modale Formeln $f(A, B)$ geben, so daß B diskussiv gültig abgelöst werden kann, sobald A als Prämisse zur Verfügung steht. "Das ist der Grund", so schreibt Jaśkowski [p. 150], "warum es in der Suche nach einer 'Logik der Diskussion' die vorrangigste Aufgabe ist, eine Funktion auszuwählen, die [...] eine Rolle spielt, welche analog zur Rolle der Implikation in gewöhnlichen Systemen ist.” Zu dieser Rolle gehört wesentlich die Modus Ponens-Regel.

Es gibt mehrere Kandidaten für eine modale Formel $f(A,B)$, welche die gewünschte Rolle übernehmen können, d.h. für welche die Sequenz

(mp) $$f(A,B), A \models_{\mathbf{D5}} B$$

gültig ist. Jaśkowski wählt als diskussive Implikation $f(A,B)$ die Formel $\Diamond A \to B$ aus und führt dafür einen neuen Junktor ein:

$$A \to_d B := \Diamond A \to B.$$

Die Modus Ponens-Eigenschaft ist für $\to_d$ garantiert, da $\Diamond(\Diamond A \to B), \Diamond A \models \Diamond B$ in **S5** gültig ist.

Ebenso können wir mit der “fehlenden” Adjunktionsregel verfahren. Wir müssen Prämissen adjungieren können, damit sie inbesondere zusammen als Antezedens für $\to_d$-Formeln zur Verfügung stehen können. Es muß also modale Formeln $g(A,B)$ geben, so daß

(adj) $$A, B \models_{\mathbf{D5}} g(A,B)$$

gültig ist. Wieder gibt es für $g(A,B)$ mehrere Kandidaten, von denen Jaśkowski die Formel $\Diamond A \wedge B$ wählt. Die diskussive Konjunktion definiert er daher so:

$$A \wedge_d B := \Diamond A \wedge B.$$

Nun wird die diskussive Adjungierbarkeit durch die in **S5** gültige Sequenz $\Diamond A, \Diamond B \models \Diamond(\Diamond A \wedge B)$ sichergestellt.

Als Resultat dieser Überlegungen dürfen wir festhalten, daß es in **D5** Junktoren gibt, die zumindest einige der für die Implikation bzw. die Konjunktion typischen Eigenschaften haben. Jaśkowski stellt zwei dieser Junktoren heraus, aber merkt sogleich an, daß auch andere die gewünschten Eigenschaften haben. Z.B. erfüllt auch $f(A,B) = \Diamond A \to \Diamond B$ die Bedingung (mp), und sowohl $g(A,B) = A \wedge \Diamond B$ als auch $g(A,B) = \Diamond A \wedge \Diamond B$ erfüllen (adj). Die Auszeichnung von $\to_d$ und $\wedge_d$ ist nicht weiter begründet und Jaśkowski scheint sie auch nur in einem beispielhaften Sinne gemeint zu haben.

Es sei nun FML^d die kleinste Menge der Formeln, die aus dem Abschluß einer gegebenen Menge atomarer Formeln unter den Junktoren in $\{\neg, \wedge, \wedge_d, \to_d\}$ entsteht. Weitere Wahrheitsfunktionen seien wie üblich aus $\{\neg, \wedge\}$ definiert. Wir übersetzen FML^d nach $\mathrm{FML}^\Diamond$ so:

Für Atome P sei $\iota(P) = P$; $\iota(\neg A) = \neg\iota(A)$; $\iota(A \wedge B) = \iota(A) \wedge \iota(B)$; $\iota(A \wedge_d B) = \Diamond\iota(A) \wedge \iota(B)$; $\iota(A \rightarrow_d B) = \Diamond\iota(A) \rightarrow \iota(B)$.

Dann ist

$$\mathbf{D5}^d = \{A \in \mathrm{FML}^d : \Diamond\iota(A) \in \mathbf{S5}\}.$$

$\mathbf{D5}^d$ (das ist etwa Jaśkowskis $\mathbf{D_2}$)[12] ist also wie **D5**, nur in einer Sprache, die um $\rightarrow_d$ und $\wedge_d$ ergänzt ist und die keine expliziten modalen Operatoren enthält.

Es ist klar, daß wir im Übergang von **D5** zu $\mathbf{D5}^d$ an Ausdrucksstärke verlieren, ohne für das Anliegen diskussiver Logik – eine logische Theorie von Diskussionen abzugeben – etwas zu gewinnen. Der Übergang wirft aber die interessante, technische Frage auf, wie $\mathbf{D5}^d$ in FML^d axiomatisiert werden kann. Ferner können diskussive Varianten weiterer Junktoren – wie der Disjunktion oder der Äquivalenz – auf verschiedene Weisen definiert werden. Schließlich können wir diskussive Logiken auf anderen modalen Systemen als **S5** aufbauen. Jaśkowskis Idee einer diskussiven Logik läßt so Raum für viele Variationen, deren Erkundung naturgemäß eine beträchtliche Menge an Literatur hervorgebracht hat.[13] Zur Frage, wie gut diskussive Logik als parakonsistente Logik ist, tragen solche Untersuchungen jedoch nichts Entscheidendes bei.

Der bestechend einfachen Definition diskussiver Folgerung auf der Basis von **S5** liegt eine ebenso bestechend einfache Idee zugrunde: Wir können den modalen Operator im Sinne eines Existenzquantors über Diskussionsbeiträge interpretieren. Im Kern überlagert **D5** damit die Folgerungsrelation von **S5** nur mit einer Relation der Möglichkeitsübertragung. Eine Sequenz $X \models_{\mathbf{D5}} A$ sagt soviel wie: In jedem Modell wird die Möglichkeit der Prämissen in X auf die Möglichkeit der Konklusion A übertragen – wobei 'Möglichkeit' im Sinne von **S5** zu verstehen ist. Möglichkeitsübertragung ist jedoch keine Wahrheitsübertragung. Wenn letztere eine notwendige Bedingung für Folgerung ist, dann ist Möglichkeitsübertragung gar keine Folgerungsbeziehung. Dieser Überlegung kann man nur dadurch entgehen, indem man eine Sequenz $X \models_{\mathbf{D5}} A$ in einem bestimmten, synkategorematischen Sinne interpretiert. Statt zu sagen "Wenn alle Aussagen in X wahr sind, dann ist auch A wahr", nehmen wir $\models_{\mathbf{D5}}$ als Auslöser für die besondere Interpretation: Wenn es wahr ist, daß A_i möglich ist (für jedes A_i in X), dann ist es auch wahr, daß A möglich ist. Oder in der diskussiven Interpretation: Wenn es wahr ist, daß jemand A_i vorgebracht hat (für jedes A_i in

---

[12] Tatsächlich ist $\mathbf{D_2}$ in [156] in einer Junktorenmenge $\{\neg, \wedge, \rightarrow_d, \leftrightarrow_d\}$ definiert, mit $A \leftrightarrow_d B := (\Diamond A \rightarrow B) \wedge (\Diamond B \rightarrow \Diamond A)$. Diese Definition von $\leftrightarrow_d$ ist sicher genauso bloß beispielhaft zu verstehen wie die von $\rightarrow_d$ und $\wedge_d$.

[13] Zum Beispiel: [217] zur Frage der Axiomatisierung, [216] über schwächere modale Grundsysteme.

X), dann ist es auch wahr, daß jemand A vorgebracht hat. Als Entwurf für eine Folgerungsrelation, die ganz allgemein Widersprüche nicht trivialisiert, taugt das jedoch nur unter der Annahme, daß Widersprüche in Theorien immer nur in der Folge unzulässiger Verschmelzungen unverträglicher Perspektiven ("Beiträge") geraten. Das trifft jedoch auf paradigmatische Fälle wie naive Wahrheits- oder Mengentheorie sicher nicht zu.

Die Einfachheit der Definition macht die in **D5** formulierte logische Diskussionstheorie zu einer wesentlich klassischen Theorie. Diskussive Logik wendet sich überhaupt nicht gegen die klassische Gleichsetzung von Inkonsistenz mit Trivialität – jedenfalls nicht, wenn die Inkonsistenz "in einem Stück", d.h. konjungiert auftritt. [14] Die Beiträge jedes Teilnehmers an einer Diskussion sind unter klassischer Folgerung abgeschlossen. Trägt ein Teilnehmer etwas vor, aus dem ein Widerspruch $A \wedge \neg A$ folgt, dann ist dieser Beitrag ohne Informationswert, denn er behauptet dann alles. Trägt ein Teilnehmer beispielsweise das naive Mengenbildungsschema vor, dann hat er gewissermaßen nichts mitgeteilt. Statt selbst eine parakonsistente Logik zu sein, bietet sich diskussive Logik vielmehr als Gegenstand einer Reform aus parakonsistenter Perspektive an. Diskussive Logik ist ein gutes Beispiel für eine Logik, die dem Buchstaben nach, jedoch nicht im beabsichtigen Sinne parakonsistent ist. Wir wenden uns nun Ansätzen zu, die das parakonsistente Anliegen in einem solchen stärkeren Sinne einzulösen versuchen. [15]

4. Der maximierende Ansatz

Zu Beginn der 60er Jahre entwickelte der brasilianische Logiker Newton da Costa eine Reihe parakonsistenter Logiken, die **C**-Systeme; siehe [56], [59], [58], [48]. Diese Systeme stießen auf großes Interesse und führten dazu, daß sich in Brasilien eine eigene und sehr aktive Schule parakonsistenter Logik herausbildete.

Da Costas Definition einer parakonsistenten Logik ist die, welche wir zu Beginn dieses Kapitels eingeführt haben, d.h.

(i) Aus A und $\neg A$ soll keine beliebige Formel B folgen.

Neben diese definitorische Bedingung stellt da Costa drei weitere Bedingungen, die eine parakonsistene Logik erfüllen sollte (vgl. [56, p. 498]):

[14] Auch andere nicht-adjunktive Ansätze versuchen im Kern klassisch vorzugehen, laufen dabei aber in ähnliche Probleme, wie sie hier für die diskussive Logik aufgezeigt werden. Siehe insbesondere die Arbeiten von Schotch und Jennings [257, 258] und das Buch [244] von Rescher und Brandom. Vgl. dazu auch Lewis [185] über "logic for equivocators". Eine kritische Diskussion dieser Ansätze findet sich in [223].

[15] Für seine vielen Hinweise zum Inhalt dieses Kapitels bin ich Max Urchs zu besonderem Dank verpflichtet.

(ii) $\neg(A \wedge \neg A)$ soll kein Theorem sein.
(iii) Die aussagenlogische Basis sollte einfach zu einer Prädikatenlogik erweitert werden können.
(iv) Die Logik sollte so weitgehend die klassische Logik **KL** enthalten, wie das mit den Bedingungen (i-iii) verträglich ist.

Im Zusammenhang mit den im Anschluß zu besprechenden dreiwertigen Logiken (vgl. p. 321) werden wir sehen, daß die zweite Bedingung nicht zwingend für jede Art parakonsistenter Logik ist. (Es wird sich jedoch gleich zeigen, daß diese Bedingung für die **C**-Systeme wesentlich ist.) Die letzten zwei Bedingungen sind natürlich etwas ungenau und lassen viele verschiedene Auslegungen zu. Die Bedingung (iii) wird uns hier nicht interessieren. Die Bedingung (iv) legt da Costa in einem sehr naheliegenden Sinne aus und setzt sie dann wörtlich in Formeln um:

> Die Logik sollte aus Formeln, die nicht widersprüchlich sind, auf klassische Weise schließen.

Für die wörtliche Umsetzung der Bedingung definiert da Costa eine Formel, die von sich selbst die Eigenschaft der Konsistenz aussagen soll:

$$A^\circ := \neg(A \wedge \neg A).$$

Die Grundidee ist jetzt sehr einfach: Formeln mit dem "Heiligenschein" $^\circ$ sollen als ungefährlich gelten; aus ihnen sollen wir klassisch schließen dürfen. Formeln ohne Heiligenschein stehen unter Verdacht (der Inkonsistenz); aus ihnen dürfen wir nur vorsichtig, d.h schwächer als klassisch schließen. Wenn wir aus einer Formel A klassisch schließen wollen, dann soll dafür die Annahme A° notwendig aber auch hinreichend sein. So soll nach der vierten Bedingung das klassische Schließen aus ungefährlichen Prämissen vollständig erhalten bleiben.

Das System $\mathbf{C}_1$. Um dieser Idee folgend eine parakonsistente Logik aufzubauen, beginnen wir mit einer Axiomatisierung des negationsfreien Fragments der klassischen Logik.

K.	$A \to (B \to A)$
S.	$A \to (B \to C) \to ((A \to B) \to (A \to C))$
$\wedge$E.	$A \wedge B \to A \qquad A \wedge B \to B$
$\wedge$I.	$A \to (B \to (A \wedge B))$
$\vee$I.	$A \to A \vee B \qquad B \to A \vee B$
$\vee$C.	$(A \to C) \to ((B \to C) \to A \vee B \to C)$
MP	$\dfrac{A \quad A \to B}{B}$

Darauf schichten wir zunächst solche Axiome für die Negation, welche die Parakonsistenz der Logik nicht in Frage stellen können.

TND. $A \vee \neg A$

DNE. $\neg\neg A \to A$

Ein klassisches Axiom fehlt in dieser Liste aus gutem Grunde, nämlich *reductio ad absurdum*,

RAA. $(A \to B) \to ((A \to \neg B) \to \neg A).$

Der gute Grund ist folgender. Die Idee einer *reductio* ist es, daß wir aus einer Annahme A etwas ableiten, was unmöglich wahr sein kann. Wenn das gelingt, dann dürfen wir auf $\neg A$ schließen. In dem Schema RAA deuten B und $\neg B$ das Unmögliche an, aus dem auf $\neg A$ geschlossen werden soll. Aber ist $B \wedge \neg B$ wirklich sicher falsch? Das Schema RAA beruht offenbar auf der Annahme, daß das so ist. Diese Einsicht, daß RAA eigentlich unter der – klassisch immer gegebenen – Annahme $\neg(B \wedge \neg B)$ steht, können wir nach da Costa in einem Schema so ausdrücken:

RAA°. $B^\circ \;\to\; ((A \to B) \to ((A \to \neg B) \to \neg A)).$

Wäre $\neg(A \wedge \neg A)$, d.h. A° ein Theorem der Logik, dann wären RAA und RAA° äquivalent. Deshalb ist es für da Costas **C**-Systeme wesentlich, daß A° nicht ableitbar ist, d.h. daß die oben genannte Bedingung (ii) erfüllt ist.

Schließlich folgt ein Schema, das festlegt, wie sich die Konsistenz von Teilformeln auf die Konsistenz von komplexeren Formeln auswirkt:

KON°. $A^\circ \wedge B^\circ \;\to\; (A \wedge B)^\circ \;\wedge\; (A \vee B)^\circ \;\wedge\; (A \to B)^\circ$

(*NB*: Die Implikation $A^\circ \to (\neg A)^\circ$ ist aus den übrigen Axiomen herleitbar.) Das vervollständigt die syntaktische Darstellung des Systems $\mathbf{C}_1$ von da Costa: die negationsfreien Axiome und die Ableitungsregel Modus Ponens, die Negationsaxiome mit RAA° statt RAA, und das Konsistenzaxiom KON°. Die Bedingungen für das Deduktionstheorem sind erfüllt, d.h. in $\mathbf{C}_1$ (wie in allen **C**-Systemen) gilt:

$$X, A \vdash B \text{ gdw } X \vdash A \to B.$$

Offensichtlich behandelt $\mathbf{C}_1$ die Negation dual zur intuitionistischen Logik **IL**: Während in dieser die TND und DNE abgelehnt, jedoch $\neg(A \wedge \neg A)$ und $A \to \neg\neg A$ beibehalten wird, so verhält es sich in $\mathbf{C}_1$ genau umgekehrt.

Der negationsfreie (positive) Teil von **IL**, $\mathbf{IL}^+$, ist jedoch in $\mathbf{C}_1$ enthalten, und $\mathbf{C}_1$ ist wiederum ein echtes Teilsystem der klassischen Logik **KL**. Also haben wir diese Beziehung:

$$\mathbf{IL}^+ \subset \mathbf{C}_1 \subset \mathbf{KL}.$$

Wenn wir zu diesen Schemata

EFQ. $\qquad A \wedge \neg A \to B$

oder das Prinzip vom ausgeschlossenen Widerspruch, $\neg(A \wedge \neg A)$, hinzufügen, dann erhalten wir eine vollständige Axiomatisierung der klassischen Logik, **KL**.

In [58, Theoreme 5-31, pp. 664-671] und [283] sind einige wichtige Formeln und Sequenzen zusammengestellt, die in $\mathbf{C}_1$ ableitbar bzw. nicht ableitbar sind. Zu den Schlüsselbeobachtungen zählt, daß keines der Kontrapositionsschemata

$$\begin{array}{ll} A \to B \vdash \neg B \to \neg A & A \to \neg B \vdash B \to \neg A \\ \neg A \to B \vdash \neg B \to A & \neg A \to \neg B \vdash B \to A \end{array}$$

gilt. Wäre es anders, dann könnten wir mit Hilfe des Axioms K EFQ ableiten. Aber $\mathbf{C}_1$ + EFQ ist einfach **KL**. Aus parakonsistenter Sicht gibt es ein direktes Argument gegen Kontraposition. Angenommen, B ist eine notwendige Bedingung für A. Wenn nun B falsch ist, dann können wir auf die Falschheit von A nur dann schließen, wenn die Falschheit von B ausschließt, daß die notwendige Bedingung B dennoch gegeben ist. D.h. wir müssen annehmen dürfen, daß B konsistent ist. Die Überlegung läßt – richtig! – erwarten, daß in $\mathbf{C}_1$ diese Sequenzen ableitbar sind:

$$\begin{array}{ll} B^\circ, A \to B \vdash \neg B \to \neg A & B^\circ, A \to \neg B \vdash B \to \neg A \\ B^\circ, \neg A \to B \vdash \neg B \to A & B^\circ, \neg A \to \neg B \vdash B \to A \end{array}$$

Es folgt dann, daß auch die Sequenz

EFQ°. $\qquad A^\circ, A, \neg A \vdash B$

in $\mathbf{C}_1$ ableitbar ist. Fügen wir zu den Axiomen von $\mathbf{C}_1$ das Schema A° hinzu (alle Formeln sind konsistent), so erhalten wir das uneingeschränkte EFQ und also die klassische Logik.

Ableitbar ist auch die Modus Ponens-Sequenz

$$A, A \to B \vdash B,$$

nicht jedoch der disjunktive Syllogismus

DS. $$A, \neg A \vee B \vdash B.$$

Die **C**-Systeme folgen also der Strategie, die Äquivalenz von $A \to B$ und $\neg A \vee B$ aufzuheben, um daraufhin MP zu bewahren und DS abzulehnen. Auch diese Entscheidung stützt ein starkes Argument. Denn der disjunktive Syllogismus lädt uns dazu ein, so zu schließen: Angenommen A-*oder*-B ist wahr. Wenn nun auch *nicht*-A wahr ist, dann kann A nicht wahr sein. Also kann die angenommene Wahrheit von A-*oder*-B nur durch die Wahrheit von B zustande kommen. Zusammengefaßt:

$$\dfrac{\dfrac{(A \vee B) \text{ wahr}}{A \text{ wahr oder } B \text{ wahr}}\,\text{I.} \qquad \dfrac{(\neg A) \text{ wahr}}{A \text{ nicht wahr}}\,\text{II.}}{B \text{ wahr}}\,\text{III.}$$

Der Schluß I ist unkontrovers.[16] Ebenso ist aus parakonsistenter Perspektive nichts gegen III einzuwenden. In dieser Perspektive nehmen wir die Möglichkeit in den Blick, daß eine Aussage als auch dessen Negation wahr sein kann. Diese Möglichkeit anzuerkennen, verpflichtet nicht zu der sehr gewagten These, daß eine Aussage zugleich wahr und nicht wahr sein kann – also etwas eine Eigenschaft zugleich haben und nicht haben kann. In III schließen wir so: A oder B hat eine Eigenschaft, A hat sie nicht; also hat B die Eigenschaft – ein guter Schluß (beim normalen Verständnis von "oder").

Gar nicht gut dagegen, ist der Schluß II. Denn II schließt einfach die Möglichkeit aus, daß A und $\neg A$ wahr sind. Aber um genau diese Möglichkeit geht es ja nun. Der disjunktive Syllogismus, d.h. die Figur

$$\dfrac{A \text{ wahr oder } B \text{ wahr} \quad (\neg A) \text{ wahr}}{B \text{ wahr}}\,\text{IV.},$$

erweist sich so als ein enthymematischer Schluß, in dem ein wichtiger Schritt – nämlich II – unterdrückt ist. Der Schluß II ist nur gerechtfertigt, d.h.

[16] Es sei denn, wir verstehen die Disjunktion in einem besonderen Sinne, der hier aber nicht gemeint ist. Eine besondere Disjunktion mit einer Wahrheitsbedingung, die den Schluß I nicht allgemein zuläßt, ist der Fissions-Junktor der Relevanzlogik; siehe Kap. VI.6.

wahrheitsübertragend, wenn wir annehmen dürfen, daß A konsistent ist. Dasselbe gilt folglich für IV. Im Gegensatz zu DS ist daher die Sequenz

DS°. $$A^\circ A, \neg A \vee B \vdash B$$

in $\mathbf{C}_1$ ableitbar.

Der Klassiker behauptet natürlich, daß die Konsistenzannahme A° immer erfüllt sei: Der Schluß IV mag – streng genommen – enthymematisch sein, aber er ist es auf risikolose Weise. Für parakonsistente Logiker ist das Risiko dagegen real, da die Annahme der Konsistenz kontingent ist. So gesehen, ist die Grundidee der C-Systeme, die Konsistenzannahmen klassischen Schließens explizit zu machen, bestechend. Wenn wir klassische Schlüsse um solche Annahmen ergänzen, dann gelten die so abgeschwächten Schlüsse in der Allgemeinheit, die wir für logisches Schließen fordern.

Die Systeme $\mathbf{C}_n$ $(0 \leq n \leq \omega)$. Eine Formel A° soll ausdrücken, daß A konsistent, d.h. daß $\neg(A \wedge \neg A)$ wahr ist. Wie steht es nun um die Aussage, daß A° konsistent ist? Ist diese Aussage, $(A^\circ)^\circ$, d.h. $\neg(\neg(A \wedge \neg A) \wedge \neg\neg(A \wedge \neg A))$, selbst konsistent? Und wie verhält sich $(A^\circ)^\circ$ (wofür wir kurz $A^{\circ\circ}$ schreiben wollen) zur Aussage, daß $A^{\circ\circ}$ konsistent ist, d.h. zu $A^{\circ\circ\circ}$ usw.? Die Bedingung, daß A konsistent sei (A°), ist offenbar schwächer als die Bedingungen, die sowohl A° als auch die Konsistenz von A° fordert ($A^\circ \wedge A^{\circ\circ}$), und letztere wiederum schwächer als die Bedingung $A^\circ \wedge A^{\circ\circ} \wedge A^{\circ\circ\circ}$, usw. Wir können also den Schluß RAA, den $\mathbf{C}_1$ unter die Bedingung A° stellt, auch unter stärkere Konsistenzbedingungen stellen – und vielleicht sollten wir das tun, um wirklich sicher unter der Bedingung möglicher Inkonsistenz zu schließen. Es ist klar, daß je stärker die Bedingung ist, die wir ins Antezedenz von RAA stellen, umso schwächer wird das Axiom und also auch die resultierende Logik. Da jede Konsistenzbedingung durch eine endliche Formel gegeben ist, kann sie durch eine stärkere Bedingung immer übertrumpft werden. Wenn uns keine endliche Bedingung genügen kann, dann sollten wir RAA ganz verbieten, wodurch offenbar eine Logik entsteht, die schwächer ist als jede, in der RAA unter irgendeiner Konsistenzbedingung steht. So entsteht die Hierarchie der $\mathbf{C}$-Systeme nach da Costa.

Für jede natürliche Zahl n (≥ 1) und Formel A definieren wir die *n-fache Konsistenz von A*, A^n, so:

$$A^1 := A^\circ \; [:= \neg(A \wedge \neg A)];$$
$$A^{n+1} := (A^n) \wedge (A^n)^\circ, \text{ für } n \geq 1.$$

Also z.B.

$$\begin{aligned} A^3 &= A^\circ \wedge A^{\circ\circ} \wedge A^{\circ\circ\circ} \\ &= \neg(A \wedge \neg A) \ \wedge \ \neg(A^\circ \wedge \neg A^\circ) \ \wedge \ \neg(A^{\circ\circ} \wedge \neg A^{\circ\circ}) \\ &= \neg(A \wedge \neg A) \ \wedge \ \neg[\neg(A \wedge \neg A) \wedge \neg\neg(A \wedge \neg A)] \ \wedge \\ &\qquad\quad \neg[\neg(A^\circ \wedge \neg A^\circ) \wedge \neg\neg(A^\circ \wedge \neg A^\circ)] \\ &= \quad \ldots \end{aligned}$$

Als Grenzfall wollen wir ferner vereinbaren, daß

$$A^0 = \top \ [:= A \vee \neg A].$$

Das schwächste der **C**-Systeme ist $\mathbf{C}_\omega$: Es ist wie $\mathbf{C}_1$, jedoch ohne RAA$^\circ$ und KON$^\circ$; oder, etwas anders beschrieben, $\mathbf{C}_\omega$ ist die Erweiterung der positiven intuitionistischen Logik $\mathbf{IL}^+$ um die "anti-intuitionistischen" Negationsaxiome TND und DNE. Nun wird für jede natürliche Zahl $n \geq 1$ die Logik $\mathbf{C}_\omega$ zur Logik $\mathbf{C}_n$ erweitert, indem das folgende Paar von Schemata hinzugenommen wird:

RAA^n	$B^n \ \rightarrow \ ((A \rightarrow B) \rightarrow ((A \rightarrow \neg B) \rightarrow \neg A))$
KON^n	$A^n \wedge B^n \ \rightarrow \ (A \wedge B)^n \ \wedge \ (A \vee B)^n \ \wedge \ (A \rightarrow B)^n$

Die klassische Logik $\mathbf{C}_0$ ist also $\mathbf{C}_\omega + \text{RAA}^0 + \text{KON}^0$, d.h. im Effekt $\mathbf{C}_\omega +$ RAA; $\mathbf{C}_1 = \mathbf{C}_\omega + \text{RAA}^1 + \text{KON}^1$, d.h. $\mathbf{C}_\omega + \text{RAA}^\circ + \text{KON}^\circ$; $\mathbf{C}_2 = \mathbf{C}_\omega + \text{RAA}^2 + \text{KON}^2 = \mathbf{C}_\omega + \text{RAA}^{\circ\circ} + \text{KON}^{\circ\circ}$ usw. Man beachte, daß in jeder Logik $\mathbf{C}_n$ das entprechende Schema

EFQ^n	$A^n, A, \neg A \vdash B$

ableitbar ist. Auf diese Beobachtung werden wir am Ende des Kapitels zurückkommen.

Wenn $[0 \neq] \ n < m$, dann $B^m \rightarrow B^n$, denn B^m ist ja lediglich eine Erweiterung von B^n um weitere Konjunkte. Daher ist RAA^m ein schwächeres Schema als RAA^n; ebenso für die KON-Schemata. Also ist $\mathbf{C}_m$ in $\mathbf{C}_n$ echt enthalten, wann immer $n < m$. Schreiben wir $L < L'$ für den Fall, daß die Menge der Theoreme der Logik L echt in der Theorem-Menge der Logik L' enthalten ist, dann können wir das auch so ausdrücken:

$$\mathbf{C}_\omega < ... < \mathbf{C}_1 < \mathbf{C}_0.$$

In dieser Hierarchie sind außer $\mathbf{C}_0$ (= **KL**) offensichtlich nur zwei Logiken von grundsätzlichem Interesse, nämlich $\mathbf{C}_1$ und $\mathbf{C}_\omega$ (alle anderen Werte für n wären willkürlich).

Die Semantik der C-Systeme. Wie der Aufbau der Axiomatik der **C**-Systeme, so baut auch deren Semantik sehr einfach auf der Semantik der klassischen Aussagenlogik auf. Im Prinzip ist alles so, wie in der klassischen Logik – bis auf die Behandlung der Negation. Eine Interpretation I der Atome in den beiden Wahrheitswerten 0 und 1 wird zu einer Interpretation $[\![-]\!]_I$ für komplexe Formeln so erweitert:

(P)	$[\![P]\!]_I = I(P)$ $(P \in \text{ATM})$
$(\wedge)$	$[\![A \wedge B]\!]_I$ gdw $[\![A]\!]_I = [\![B]\!]_I = 1$
$(\vee)$	$[\![A \vee B]\!]_I$ gdw $[\![A]\!]_I = 1$ oder $[\![B]\!]_I = 1$
$(\rightarrow)$	$[\![A \rightarrow B]\!]_I$ gdw $[\![A]\!]_I = 0$ oder $[\![B]\!]_I = 1$

Das sind die klassischen Interpretationsbedingungen. Für die Negation gibt es keine wahrheitsfunktionale Bedingung, d.h. eine, die den Wert von $\neg A$ vollständig aus dem Wert für A bestimmt. Eine Interpretation wird in ihrer Freiheit, Werte über A und $\neg A$ zu verteilen, nur durch eine Auswahl (je nach dem zu modellierenden System) aus den folgenden Bedingungen eingeschränkt. (Wir lassen jetzt den Bewertungsindex I zur Interpretation fort.)

(A) Wenn $[\![A]\!] = 0$, dann $[\![\neg A]\!] = 1$.
(B) Wenn $[\![\neg\neg A]\!] = 1$, dann $[\![A]\!] = 1$.
(C^1) Wenn $[\![A]\!] = [\![\neg A]\!] = 1$, dann $[\![A^\circ]\!] = 0$.
(C^n) Wenn $[\![A^{n-1}]\!] = [\![(\neg A)^{n-1}]\!] = 1$, dann $[\![A^n]\!] = 0$.
(D^1) Wenn $[\![A^\circ]\!] = [\![B^\circ]\!] = 1$,
dann $[\![(\neg A)^\circ]\!] = [\![(A \wedge B)^\circ]\!] = [\![(A \vee B)^\circ]\!] = [\![(A \rightarrow B)^\circ]\!] = 1$.
(D^n) Wenn $[\![A^n]\!] = [\![B^n]\!] = 1$,
dann $[\![(\neg A)^n]\!] = [\![(A \wedge B)^n]\!] = [\![(A \vee B)^n]\!] = [\![(A \rightarrow B)^n]\!] = 1$.
(Z) Wenn $[\![\neg A]\!] = 1$, dann $[\![A]\!] = 0$.[17]

Das sind Bedingungen, die unmittelbar ersichtlich die oben eingeführten Schemata spiegeln, welche die verschiedenen **C**-Systeme axiomatisch kennzeichnen. Insofern bietet diese Semantik keinen unabhängigen und schon gar keinen rechtfertigenden Zugang zu den Systemen. Sie hat jedoch durchaus eine technische Berechtigung, indem sie schnelle Verifikationen und Falsifikationen erlaubt.

---

[17] Neben (A-D) sind weitere Bedingungen denkbar, die zu verschiedenen parakonsistenten Logiken führen. So können durch solche Bedingungen zum Beispiel auch die DeMorgan-Gesetze durchgesetzt werden. Die Bedingung (Z) ist jedenfalls die "letzte" Bedingung, die parakonsistente Logiker in Kraft sehen wollen: Sie generiert, im Verein mit (A), die klassische Logik.

Unter einem *Modell* wollen wir jetzt eine Interpretation $[\![-]\!]_I$ verstehen, welche die oben genannten Bedingungen für Atome und die zweistelligen Junktoren erfüllt, sowie die Bedingungen (A) und (B). Eine Formel ist *gültig* in einer Klasse von Modellen, wenn Sie in allen Modellen dieser Klasse den Wert 1 erhält.

SATZ 4. (Richtigkeit und Vollständigkeit für die C-Systeme)

1. *Eine Formel A ist genau dann in $\mathbf{C}_\omega$ ableitbar, wenn A in allen Modellen gilt.*
2. *Für $n \geq 1$ ist eine Formel A genau dann in $\mathbf{C}_n$ ableitbar, wenn A in allen Modellen gilt, welche die Bedingungen* (C^n) *und* (D^n) *erfüllen.*
3. *Eine Formel A ist genau dann in $\mathbf{C}_0$ $(= \mathbf{KL})$ ableitbar, wenn A in allen Modellen gilt, welche die Bedingung* (Z) *erfüllen.*

Die Richtigkeit der Systeme wird durch die übliche Induktion über die Länge eines Beweises gezeigt. Auch für die Richtung der Vollständigkeit (rechts nach links), können wir im Prinzip so wie in der klassischen Logik vorgehen, indem wir uns des Lindenbaum-Lemmas bedienen (siehe dazu das Einleitungskapitel). Statt maximal konsistenter Mengen betrachten wir nun aber maximal nicht-triviale Mengen. Da wir im folgenden den Vollständigkeitsbeweis nur am System $\mathbf{C}_1$ illustrieren (d.i. der Basisfall der Behaupung 2 des Satzes) ist maximale Nicht-Trivialität stets im Sinne des Abschlusses Cn unter $\mathbf{C}_1$-Konsequenz zu verstehen.

- Eine Formelmenge X ist genau dann *trivial* (im Sinne von $\mathbf{C}_1$), wenn $\mathrm{Cn}(X) = \mathrm{FML}$; X ist *maximal nicht-trivial*, wenn $\mathrm{Cn}(X) \neq \mathrm{FML}$ und für alle $A \notin X$, $\mathrm{Cn}(X \cup \{A\}) = \mathrm{FML}$.

Im folgenden Lemma, das einfach zu beweisen ist, sollten unschwer die allgemeinen Bedingungen für ein Modell (in 1–5), sowie die besonderen Bedingungen (C^1) und (D^1) (in 6–7) erkennbar sein.

LEMMA 5. *Für jede (unter* Cn *in* $\mathbf{C}_1$*) maximal nicht-triviale Menge X gilt:*

0. *$A \in X$ gdw $X \vdash A$ in $\mathbf{C}_1$;*
1. *$A \wedge B \in X$ gdw $A \in X$ und $B \in X$;*
2. *$A \vee B \in X$ gdw $A \in X$ oder $B \in X$;*
3. *$A \rightarrow B \in X$ gdw $A \notin X$ oder $B \in X$;*
4. *Wenn $A \notin X$, dann $\neg A \in X$;*
5. *Wenn $\neg\neg A \in X$, dann $A \in X$;*
6. *Wenn $A, \neg A \in X$, dann $A^\circ \notin X$;*
7. *Wenn $A^\circ, B^\circ \in X$, dann $(\neg A)^\circ, (A \wedge B)^\circ, (A \vee B)^\circ, (A \rightarrow B)^\circ \in X$.*

LEMMA 6. *Jede (unter* Cn *in* $\mathbf{C}_1$*) maximal nicht-triviale Menge hat ein Modell, welches die Bedingungen* (C^1) *und* (D^1) *erfüllt.*

BEWEIS. Gegeben eine maximal nicht-triviale Menge X, definieren wir eine Funktion $[\![-]\!]_X$ so: Für jede Formel A sei $[\![A]\!]_X = 1$, falls $A \in X$; $[\![A]\!]_X = 0$ anderenfalls. Aufgrund des vorigen Lemmas können wir dann verifizieren, daß jede solche Funktion $[\![-]\!]_X$ eine Interpretation unter den Bedingungen (A) und (B), also ein Modell ist. So generierte Modelle erfüllen ferner die Bedingungen (C^1) und (D^1). ∎

Das nächste Lemma wird kaum anders bewiesen als im Fall der klassischen Logik (also "$\mathbf{C}_0$"). Wieder sei hier maximale Nicht-Trivilität durch Abschluß unter Cn in $\mathbf{C}_1$ bestimmt.

LEMMA 7. (Nach Lindenbaum) *Jede nicht-triviale Menge läßt sich zu einer maximal nicht-trivialen Menge erweitern.*

Das Argument für die Vollständigkeit von $\mathbf{C}_1$ führen wir nun zuende, indem wir eine beliebige Formel F betrachten, die in $\mathbf{C}_1$ nicht beweisbar ist. (Wir wollen ein Modell finden mit $[\![F]\!] = 0$.) Da F nicht ableitbar ist, ist die Menge $\{\neg F, F^\circ\}$ nicht trivial. Wir erweitern diese Menge nach dem Lindenbaum-Lemma zu einer maximal nicht-trivialen Menge F^*. Diese Menge hat ein Modell, welches (C^1) erfüllt. Die Bedingung besagt, daß wenn $[\![\neg F]\!] = [\![F^\circ]\!] = 1$, dann $[\![F]\!] = 0$. Das aus F^* generierte Modell erfüllt offensichtlich den Wenn-Teil $[\![\neg F]\!] = [\![F^\circ]\!] = 1$. Also $[\![F]\!] = 0$, wie gewünscht. ∎

Damit ist Satz 4.2 bewiesen. Für ein beliebig gewähltes System $\mathbf{C}_n$ ($1 < n$) verfährt die Argumentation völlig analog. Es muß in den Lemmata nur n für 1 eingesetzt und die (C^n)- und (D^n)-Schemata aufgerufen werden. Im Falle von $\mathbf{C}_\omega$ entfällt die Verifikation von besonderen Modellbedingungen. Für $\mathbf{C}_0$ (klassische Logik) koinzidieren maximal nicht-triviale Mengen mit den vertrauten maximal konsistenten Mengen, in denen $\neg A \in X$ gdw $A \notin X$ (Bedingung Z) gilt.

Substitutionsfehler. Fragen wir zunächst, wie die **C**-Systeme im Lichte der Kriterien (i-iv) von da Costa dastehen. Die ersten zwei Bedingungen sind einfach zu verifizieren ($1 \leq n \leq \omega$):

(i) In $\mathbf{C}_n$ gilt nicht: $A, \neg A \vdash B$.
(ii) In $\mathbf{C}_n$ ist $A \wedge \neg A$ kein Theorem.

Die dritte Bedingung – der einfache Aufstieg zur ersten Stufe – interessiert uns hier nicht. Die vierte Bedingung, nach der die C-Logiken die Stärke der klassischen Logik weitmöglichst approximieren sollten, ist in den $\mathbf{C}_n$ Logiken (n endlich) auf folgende Weise realisiert:

(iv) $X^n, X \vdash A$ in $\mathbf{C}_n$ gdw $X \vdash A$ in **KL**.

Mit anderen Worten, unter der Annahme, daß die Prämissen einer Sequenz konsistent sind, wird wie in der klassischen Logik geschlossen. Aber wie stark kann in den **C**-Systemen geschlossen werden, wenn die Schlüsse nicht unter Konsistenzannahmen stehen, d.h. in den Fällen, die uns hier besonders interessieren?

Es stellt sich heraus, daß die **C**-Systeme einige merkwürdige Barrieren aufbauen, die aus parakonsistenten Absichten nicht begründet werden können. Vielleicht am auffälligsten und bedauernswertesten ist der Umstand, daß das Austauschprinzip versagt: Wenn A und A' logisch äquivalent (gegenseitig ableitbar) sind, dann sollten es auch $B[A]$ und $B[A']$ sein; in Form einer Beweisregel:

SUB. $$\frac{A \equiv A'}{B[A] \equiv B[A']},$$

wobei $B[A]$ eine Formel andeute, in der A vorkommt, und $B[A']$ aus B durch Substituierung beliebiger Vorkommen von A durch A' ensteht.

Wir erwarten von einer Logik, daß sie Formeln nicht bloß als syntaktische Objekte, sondern als Informationsträger auffaßt. Nun können wir selbstverständlich Sprachen betrachten, in denen logische Äquivalenz kein hinreichendes Identitätskriterium für Information abgibt. Sprachen, in denen doxastische Kontexte gebildet werden können (soundso glaubt, daß ...), mögen von dieser Art sein. Aber im Vokabular der **C**-Systeme gibt es keine Ausdrücke, die *solche* Substitutionsfehler erwarten ließen. Wenn zwei Sätze einer **C**-Sprache logisch dieselbe Information enthalten, dann sollten sie in allen Kontexten logisch gleich behandelt werden. In den **C**-Systemen ist das jedoch nicht der Fall.

Zunächst die gute Nachricht: Da $\mathbf{IL}^+ \subseteq \mathbf{C}_1$, und $\mathbf{IL}^+$ unter SUB abgeschlossen ist, so ist auch der negationsfreie Teil von $\mathbf{C}_1$ unter SUB abgeschlossen. Es ist also die Negation, welche Kontexte hervorbringt, die gegen die Ersetzbarkeit logisch gleicher Teilformeln abgeschirmt sind. Beispielsweise ist in allen **C**-Systemen die Konjunktion idempotent und kommutativ,

$$A \equiv A \wedge A \qquad A \wedge B \equiv B \wedge A.$$

Aber im Skopus der Negation ist das in keinem der Systeme allgemein der Fall. Es gilt

nicht : $$\neg A \equiv \neg(A \wedge A) \qquad \neg(A \wedge \neg A) \equiv \neg(\neg A \wedge A).$$

Aus semantischer Perspektive liegt das daran, daß der Junktor $\neg$ keine Wahrheitsfunktion ist. So erlaubt uns die Bedingung (A) zwar auf den Wert von $\neg A$ zu schließen, wenn $[\![A]\!] = 0$ ($\neg A$ hat dann den Wert 1); aber wenn $[\![A]\!] = 1$, dann sagen die Negations-Bedingungen für die nicht-klassischen **C**-Systeme nichts allgemein über den Wert von $\neg A$ aus. Deshalb gibt es falsifizierende Interpretationen für jede Richtung der zwei Ableitungsäquivalenzen.

> Wir betrachten eine atomare Instanz der ersten Äquivalenz, $\neg P \equiv \neg(P \wedge P)$. Wenn P den Wert 1 erhält, dann hindert uns nichts daran, $\neg P$ den Wert 1 und $\neg(P \wedge P)$ den Wert 0 (oder umgekehrt) zu geben. — Ähnlich verhält es sich mit der zweiten Äquivalenz. Wir betrachten eine Instanz mit einem Atom P, das den Wert 1 hat. Dann erlauben die Bedingungen $[\![P \wedge \neg P]\!] = [\![\neg P \wedge P]\!] = 1$, während wir den Negationen der Formeln $P \wedge \neg P$ und $\neg P \wedge P$ verschiedene Werte zuweisen können. (Urbas [283, Theorem 1] zeigt dasselbe Resultat anhand von fünfwertigen Matrizen.)

In der schwächsten Logik $\mathbf{C}_\omega$ versagt SUB sogar vollständig in Negationskontexten: $\neg A \equiv \neg B$, nur wenn $A = B$ [283, Theorem 2]!

In einer parakonsistenten Logik gibt es keinen wirklich guten Grund, warum es nicht (logisch) einerlei sein soll, ob man A oder $A \wedge A$ negiert. Und auch im Skopus der Negation sollte man Konjunkte ohne Bedenken vertauschen können. Daher ist überhaupt nicht einzusehen, warum nur $\neg(A \wedge \neg A)$, nicht aber $\neg(\neg A \wedge A)$ die Konsistenz von A, A°, ausdrücken soll. Diesen Defekt können wir nicht einfach dadurch beheben, daß wir SUB als Regel zu den Systemen hinzunehmen. Denn, wie Urbas [283] gezeigt hat, werden alle Systeme $\mathbf{C}_n$ (endliches n) zur vollen klassischen Logik, sobald wir die Menge der Theoreme unter SUB abschließen.

Endliche Trivialisierbarkeit. Gegen die $\mathbf{C}_n$-Systeme ($0 < n < \omega$) läßt sich ein weiterer, grundsätzlicher Einwand erheben. Um diesen Einwand zu verstehen, erinnern wir uns daran, was parakonsistente Logiker an der klassischen Logik beklagen. Auch wenn wir uns um eine sorgfältige Bildung unserer Überzeugungen bemühen, so kann es dennoch geschehen, daß durch das Hinzufügen einer *einzigen* weiteren Überzeugung A zu unserer bisherigen Überzeugungsmenge K unsere Überzeugungen trivial werden, d.h. $\mathrm{Cn}(K \cup \{A\}) = \mathrm{FML}$. Das geschieht immer dann, wenn K bereits $\neg A$ enthält und also der Konsequenzoperator auf eine Menge angewandt wird, aus der ein Satz der Form $A \wedge \neg A$ ableitbar ist. Etwas allgemeiner wollen wir sagen:

- Eine Logik L ist *endlich trivialisierbar*, wenn es endliche Formelmengen X gibt, deren Abschluß unter Cn_L trivial ist, also $\mathrm{Cn}_L(X) = \mathrm{FML}$; in einem solchen Fall sagen wir, daß L die Menge X *trivialisiert*.

Jede endliche Menge mit einer Teilmenge der Form $\{A, \neg A\}$ belegt die endliche Trivialisierbarkeit von **KL**. Jede Logik L, die endlich trivialisierbar ist, droht mit dem gerade beschriebenen Szenario: Eine bisher "gute" (d.h. zumindest nicht-triviale) Theorie K schlägt in die Menge aller Formeln um, sobald K "unglücklich" um eine Formel erweitert und unter Cn_L abgeschlossen wird. Wenn die endliche Trivialisierbarkeit der klassischen Logik der eigentliche Stein des Anstoßes ist, dann sollte eine parakonsistente Logik diese Eigenschaft nicht mit **KL** teilen.

Jedes der $\mathbf{C}_n$-Systeme ist aber endlich trivialisierbar! Das folgt unmittelbar aus der Ableitbarkeit der Sequenz EFQ^n, $A, \neg A, A^n \vdash B$. Jede Menge, die $\{A, \neg A, A^\circ\}$ enthält, wird durch $\mathbf{C}_1$ trivialisiert; und allgemein trivialisiert $\mathbf{C}_n$ jede Menge, die $\{A, \neg A, A^n\}$ enthält. Die $\mathbf{C}_n$-Systeme sind zwar parakonsistent im Wortsinn der Definition, verfehlen aber die eigentlich gemeinte Absicht genauso wie die klassische Logik. Das liegt an der Art, wie die vierte Bedingung in diesen Systemen umgesetzt wird.

(iv) Die Logik sollte so weitgehend die klassische Logik **KL** enthalten wie das mit den Bedingungen (i-iii) verträglich ist.

In den **C**-Systemen soll "im normalen Fall" klassisch geschlossen werden dürfen. Der normale Fall ist, daß unsere Prämissen konsistent sind. Die Annahme, daß wir uns im normalen Fall befinden, wird durch eine Formel A° ausgedrückt. Aber das Hinzufügen von A° zu unseren Annahmen, d.h. die *Behauptung*, daß $\neg(A \wedge \neg A)$, garantiert natürlich nicht, daß wir uns *tatsächlich* im normalen Fall befinden, in dem wir gefahrlos klassisch schließen können.

Wozu endliche Trivialisierbarkeit konkret führen kann, wollen wir am Beispiel einer Wahrheitstheorie (siehe p. 290) als Erweiterung von $\mathbf{C}_1$ illustrieren.[18] Zunächst stellen wir eine Beobachtung voran:

- Es sei $-A := \neg A \wedge A^\circ$ (starke Negation). Dann sind die Theoreme von $\mathbf{C}_1$ im $\{-, \wedge, \vee, \rightarrow\}$-Fragment der Sprache genau die Theoreme der klassischen Logik (wenn wir die klassische Negation mit $-$ bezeichnen). (Für einen Beweis genügt es, die klassischen Axiome mit $-$ als Negation in $\mathbf{C}_1$ abzuleiten. Alternativ genügt auch die Beobachtung, daß $[\![A]\!] = 0$ gdw $[\![-A]\!] = 1$ in den Modellen für $\mathbf{C}_1$ gilt.)

Sei $\mathbf{T}_1$ die Erweiterung von $\mathbf{C}_1$ um das T-Schema in einer Sprache mit genügend Ressourcen, um auf ihre eigenen Sätze Bezug nehmen zu können. Wenn wir den Lügnersatz L ("$\ulcorner L \urcorner$ ist nicht wahr") in $\mathbf{T}_1$ betrachten, so folgt die Äquivalenz $L \leftrightarrow \neg L$ und dann auf bekannte Weise, daß $L \wedge \neg L \in \mathbf{T}_1$.

---

[18] Die Illustration stammt von da Costa selbst; siehe [56].

Da aber EFQ nicht gilt, so ist $\mathbf{T}_1$ zwar widersprüchlich, jedoch nicht trivial. So weit, so gut.

Nun betrachten wir den starken Lügnersatz K = "$\ulcorner K \urcorner$ ist stark falsch"; oder direkter:

$$K: \qquad \ulcorner -K \urcorner \text{ ist wahr.}$$

Dann können wir $K \leftrightarrow -K$ in $\mathbf{T}_1$ ableiten. Aufgrund der Beobachtung sind $(A \leftrightarrow -A) \to A \wedge -A$ und $A \wedge -A \to B$ in $\mathbf{T}_1$; also ist die Wahrheitstheorie $\mathbf{T}_1$ trivial. Das gleiche Argument können wir für jede Wahrheitstheorie $\mathbf{T}_n$ auf der Basis eines der Systeme $\mathbf{C}_n$ wiederholen, wenn wir die Definition der starken Negation entsprechend anpassen.

Das System $\mathbf{C}_\omega$ verzichtet ganz auf den Versuch, die Bedingung (iv) mithilfe von Normalitätsannahmen umzusetzen. Das System ist *nicht* endlich trivialisierbar und also in einem substanziellen Sinne parakonsistent. Semantisch wird es einfach durch die Abschwächung der klassischen Äquivalenz $[\![\neg A]\!] = 1$ gdw $[\![A]\!] = 0$ zu nur einer Richtung (von rechts nach links) und der Bedingung, wenn $[\![\neg\neg A]\!] = 1$ dann $[\![A]\!] = 1$, festgelegt. Die erste Bedingung entspricht dem Vollständigkeitsprinzip TND, die zweite dem Schema DNE. $\mathbf{C}_\omega$ leidet, wie erwähnt, an einem Totalausfall der SUB-Regel: *Keine* syntaktisch verschiedenen Formeln lassen sich im Skopus der Negation füreinander einsetzen. Im Gegensatz zu den $\mathbf{C}_n$-Systemen läßt sich dieser Fehler jedoch beheben. Der Abschluß von $\mathbf{C}_\omega$ unter

$$\frac{A \to B}{\neg B \to \neg A} \quad \text{oder der schwächeren Regel} \quad \frac{A \leftrightarrow B}{\neg A \to \neg B}$$

garantiert den Abschluß unter SUB, bleibt dabei unterhalb der Schwelle der klassischen Logik und ist weiterhin parakonsistent im starken Sinne, der die endliche Trivialisierbarkeit ausschließt (siehe Urbas [283]). Jedoch beschneidet die Negationstheorie von $\mathbf{C}_\omega$ (oder einer der zwei Erweiterungen) die der klassischen Logik deutlich stärker, als es das Ziel der Parakonsistenz erfordert. So gibt es in ihr keine Kontrapositions- oder *reductio*-Prinzipien. In jedem Anwendungsfall ensteht daher die Frage, ob der Abschluß bereichsspezifischer Axiome unter $\mathbf{C}_\omega$ eine genügend starke Theorie hervorbringt. Einige Beispiele werden in [58] beschrieben.

5. Der wahrheitsfunktionale Ansatz

Parakonsistente Logiker weisen darauf hin, daß manche Sätze *wahr*, *falsch* oder *beides*, d.h. wahr-und-falsch, sein können. Nichts naheliegender also, als parakonsistente Logik als eine Logik dreier Wahrheitswerte aufzufassen: Wahr (1), Falsch (1) und Beides (b).

Jede wahrheitsfunktionale Logik bestimmt den Wahrheitswert eines zusammengesetzten Satzes allein als eine Funktion der Wahrheitswerte seiner Teilsätze. Das garantiert, daß gleichwertige Sätze überall einander ersetzen können. Die **C**-Systeme des letzten Abschnitts sind nicht wahrheitsfunktional und haben deshalb, wie wir gesehen haben, inbesondere Schwierigkeiten mit der Substitutionsregel SUB. In einer wahrheitsfunktionalen Logik über drei Werte (ab jetzt kurz: dreiwertige Logik) können solche Schwierigkeiten erst gar nicht entstehen.

Eine dreiwertige parakonsistente Logik hat den Vorteil, besonders einfach darstellbar und entscheidbar zu sein. Die Junktoren der Sprache werden als Wahrheitsfunktionen interpretiert, und diese wiederum in den vertrauten Wahrheitstafeln dargestellt. Wenn $*$ ein zweistelliger Junktor ist, dann gibt es 3^2 Möglichkeiten, wie die drei Werte über die zwei Argumente A und B der Funktion verteilt sein können. Für jede der neun Möglichkeiten muß dann eine Festsetzung für den Wert von $A * B$ getroffen werden. Die Wahrheitswertverteilungen über A und B können wir als Koordinatenpaare darstellen und den Wert für $A * B$ an den jeweiligen Kreuzungspunkten eintragen. So könnte z.B. diese Tafel enstehen:

$*$	1	b	0
1	1	b	0
b	b	b	0
0	0	0	0

Man beachte, daß in dieser Tafel die "klassischen Paare" (1,1), (1,0), (0,1) und (0,0) auf die "klassischen Werte" 0 oder 1 abgebildet werden. Dahinter steht die Idee, daß wir in einer konsistenten Situation den Wert b ignorieren dürfen, die Interpretation von $A*B$ also zweiwertig ist. Im Falle dieser Tafel sehen wir, daß der zweiwertige Teil der Tafel mit der klassischen Tafel für die Konjunktion übereinstimmt. So können wir da Costas Bedingung (iv) recht einfach umsetzen: In konsistenten Situationen sollen die Junktoren wie in der klassischen Logik interpretiert werden. Nur, wenn A oder B inkonsistent ist – den Wert b erhält – müssen wir uns überlegen, welchen Wert $A*B$ haben soll. In einer dreiwertigen parakonsistenten Logik müssen wir also nicht infrage stellen, wie die klassische Logik die Junktoren interpretiert –

unter der Annahme, daß keine Formel sowohl wahr als auch falsch sein kann. Vielmehr heben wir die Annahme auf und fragen, wie die klassische Interpretation nun unter Einbeziehung des Wertes b plausibel erweitert werden kann. Die Agenda ist also klar: Wir gehen von der klassischen Verteilung der Werte 0 und 1 für die wichtigsten Junktoren $\neg, \wedge, \vee$ und $\rightarrow$ aus und vervollständigen die Tafeln für den Wert b.

Die Logik LP. Es stellt sich schnell heraus, daß es gar nicht sehr viele plausible Möglichkeiten gibt. Die sogenannten starken Kleene-Tafeln sind besonders einleuchtend:[19]

Die starken Kleene-Tafeln

$\neg$	
1	0
b	b
0	1

$\wedge$	1	b	0
1	1	b	0
b	b	b	0
0	0	0	0

$\vee$	1	b	0
1	1	1	1
b	1	b	b
0	1	b	0

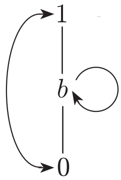
$\rightarrow$	1	b	0
1	1	b	0
b	1	b	b
0	1	1	1

Die Eckwerte sind, wie schon erwähnt, durch die klassischen Tafeln vorgegeben. Wenig plausibel wäre es, davon auszugehen, daß ein mit b bewerteter Teilsatz jeden Satz, von dem er ein Teil ist, mit b infiziert. (Das wären die sogenannten schwachen Kleene-Tafeln, in denen alle b-Zeilen und -Spalten durchgängig mit b gefüllt werden.)

- So ist eine Konjunktion mit einem falschen Konjunkt sicher falsch. Also $0 \wedge b = b \wedge 0 = 0$.
- Ebenso gilt, daß für die Wahrheit einer Disjunktion die Wahrheit eines der Disjunkte reicht. Also $1 \vee b = b \vee 1 = 1$.
- Für die Implikation folgen die starken Kleene-Tafeln der klassischen Vorgabe, daß $A \rightarrow B$ gleichbedeutend mit $\neg A \vee B$ sein soll. Die Tafel für $\rightarrow$ errechnet sich also aus den Tafeln für $\neg$ und $\vee$ und ist in diesem Sinne redundant.

Wir können die Kleene-Tafeln auch in einem Hasse-Diagramm zusammenfassen:

[19] Siehe [161], ch. 12, §64. In Kleenes Interpretation soll der dritte Wert die Situation vorstellen, in der eine Aussage weder wahr noch falsch, d.h. in diesem Sinne "unbestimmt" (Kleene: "undefined") ist. Die Tafeln heißen bei Kleene "stark", weil sie die 2-wertigen Tafeln in den b-Reihen und -Spalten in einem Sinne, den wir gleich erläutern werden, maximal um die Werte 0 und 1 ergänzen.



Die Pfeile deuten die Abbildung eines Wertes unter der Negation an. Konjunktion “geht nach unten”, Disjunktion “geht nach oben”. Damit ist folgendes gemeint: $x \wedge y$ ist immer der niedrigere (“untere”) und $x \vee y$ ist immer der höhere (“obere”) der beiden Werte x und y. So werden in den $\wedge$- und $\vee$-Tafeln die Eckwerte 0 und 1 jeweils so maximiert, wie das auch in der zweiwertigen Logik geschieht: $\wedge$ ist die Minimum- und $\vee$ ist die Maximum-Funktion.

Die Tafeln machen Aussagen über den Wahrheitswert eines Satzes unter Einschluß der Möglichkeit, daß Sätze wahr *und* falsch sein können. Was ist nun ein logisch wahrer Satz und wann folgt ein Satz aus Prämissen?

Um diese Frage zu beantworten, hilft es, sich den Wert b (“Beides”) als zusammengesetzt vorzustellen – etwa durch die Menge $\{w, f\}$ dargestellt. Den Wert rein wahrer Sätze (1) könnten wir dann durch die Menge $\{w\}$ und den Wert rein falscher Sätze (0) durch $\{f\}$ repräsentieren. Das macht deutlich, daß auch Sätze mit dem Wert b wahr sind – jedenfalls ein bißchen wahr.

Ein Satz ist logisch wahr, wenn er unter jeder Interpretation wahr ist. Das kann nach dem soeben Gesagtem zweierlei bedeuten: A ist genau dann logisch wahr, wenn A unter jeder Interpretation wahr ($\{w\}$ oder $\{w, f\}$) oder, strikter, rein wahr ($\{w\}$) ist. D.h.

(a) für alle $I: \ I(A) \in \{1, b\}$; oder
(b) für alle $I: \ I(A) = 1$.

Analoge Optionen gibt es auch für die Definition einer logischen Folgerung. Wenn X eine Formelmenge ist, dann sei $I(X) = 1$ bzw. $I(X) \in \{1, b\}$ kurz für: $\forall B \in X$, $I(B) = 1$ bzw. $I(B) \in \{1, b\}$. Eine Formel A folgt genau dann aus einer Prämissenmenge X, wenn

(a) für alle I, wenn $I(X) \in \{1, b\}$, dann $I(A) \in \{1, b\}$; oder
(b) für alle I, wenn $I(X) = 1$, dann $I(A) = 1$.

In der Terminologie der mehrwertigen Logik kann man die zur Wahl stehenden Möglichkeiten auch so beschreiben: In der einen, (a), sind die Werte 1 und b *designiert*; in der anderen, (b), ist nur der Wert 1 designiert. (Die desiginierten Werte sind gewissermaßen die “guten” Werte.)

Welche der beiden Optionen eignet sich nun für die Definition des Begriffs einer gültigen Formel bzw. Folgerung in einer parakonsistenten Logik? Ausgehend von der Definition einer parakonsistenten Logik betrachten wir zuerst den Folgerungsbegriff. Wir müsssen vermeiden, daß die Sequenz

sEFQ. $$A, \neg A \vdash B$$

gilt.[20] Die Folgerung gilt aber – auf leere Weise – unter der Definition (b). Denn die Voraussetzung $I(\{A, \neg A\}) = 1$ ist in den Tafeln nicht zu erfüllen. Also trifft es zu, daß wenn $I(\{A, \neg A\}) = 1$, dann $I(B) = 1$.

Legen wir dagegen die Definition (a) zugrunde, dann können wir z.B. der Menge $\{A, \neg A\}$ den Wert b geben, wenn $I(A) = b$ ist. Das erzwingt ersichtlich nicht den Wert 1 oder b für beliebige Formeln B. So zeigt eine Interpretation zweier Atome P und Q mit $I(P) = b$ und $I(Q) = 0$, daß sEFQ nicht gültig sein kann.

Im Anschluß an Graham Priest [221] nennen wir die durch (a) festgelegten Mengen gültiger Formeln und Folgerungen die Logik **LP** (*Logic of Paradox*). **LP** ist, wie wir gerade gesehen haben, parakonsistent. Sie ist dies offensichtlich in einem deutlich gehaltvolleren Sinne als die diskussive Logik. Die Möglichkeit widersprüchlicher Formeln geht **LP** direkt und allgemein an, indem diese Möglichkeit durch einen dritten Wert neben den klassischen Werten 0 und 1 dargestellt wird. Der Rest der Theorie erhält sich die Einfachheit jeder wahrheitsfunktionalen Logik.

Anmerkung: Die Kleene-Logik **L3**. Wenn wir, wie in **LP**, Formeln nach den starken Kleene-Tafeln auswerten, jedoch für die Definition der Folgerungsrelation nur den Wert 1 desiginieren, d.h. der Option (b) folgen, dann haben wir damit die sogenannte Kleene-Logik **L3** definiert. In dieser Logik ist die Interpretation des dritten Wertes als "unbestimmt", d.h. weder wahr noch falsch, sinnvoll; siehe [161]. Wir haben hier keinen Anlaß, uns in die Eigenschaften von **L3** zu vertiefen. Die Logik **L3** wird uns im Kapitel VI über Relevanzlogik als, gewissermaßen, eine Hälfte der grundlegenden DeMorgan-Logik **DML** wiederbegegnen. (*Ende der Anmerkung.*)

Wir haben gesehen, daß es **LP** auf sehr einfache Weise gelingt, Inkonsistenz von Trivialität zu separieren. Aber wie erfolgreich ist **LP** als eine *logische* Theorie?

Zunächst dürfen wir festhalten, daß **LP** erstaunlich klassisch ist. Wir wollen zwei Formeln *stark äquivalent* nennen, wenn sie unter jeder Bewertung

[20] Wir schreiben hier Sequenzen weiter mit dem einfachen Drehkreuz $\vdash$, obwohl wir die Folgerungsrelation für **LP** semantisch definiert haben und deshalb $\models$, wie schon im Abschnitt über diskussive Logik, die konsequentere Notation wäre.

denselben Wert, d.h. dieselben Wahrheitstafeln haben (Notation: $A \rightleftharpoons B$). Die folgenden starken Äquivalenzen gelten in **LP**:

$$A \rightleftharpoons \neg\neg A$$
$$A \to B \rightleftharpoons \neg B \to \neg A$$
$$\neg A \vee \neg B \rightleftharpoons \neg(A \wedge B) \qquad \neg A \wedge \neg B \rightleftharpoons \neg(A \vee B)$$
$$A \wedge B \rightleftharpoons \neg(\neg A \vee \neg B) \qquad A \vee B \rightleftharpoons \neg(\neg A \wedge \neg B)$$
$$A \to B \rightleftharpoons \neg A \vee B$$

Es bleiben also in **LP** die elementaren Beziehungen zwischen den Wahrheitsfunktionen und die daraus folgenden, wechselseitigen Definierbarkeiten, die wir aus der klassischen Logik kennen, erhalten. Aus den Äquivalenzen geht auch hervor, daß **LP** den Negationsjunktor deutlich stärker an seine Boole'sche Bedeutung heranführt, als das in den **C**-Systemen der Fall ist. Die Austauschregel SUB ist kein Problem für **LP**. Es wird wie in der klassischen Logik durch eine einfache Induktion über den Formelaufbau bewiesen, ist aber schon aus dem wahrheitsfunktionalen Charakter der Junktoren evident. Überhaupt unterscheiden sich **LP** und die klassische Logik **KL** nicht in ihren Antworten auf die Frage, welche Formeln gültig sind, sondern nur in der Beurteilung von Folgerungen.

Beobachtung 8. *Eine Formel A ist genau dann gültig in* **LP**, *wenn A gültig in* **KL** *ist.*

Beweis. Von rechts nach links ist das trivial. Für die andere Richtung betrachten wir **2**, die Menge der klassischen zweiwertigen und **3** die der dreiwertigen Abbildungen von Formeln auf Wahrheitswerte; ferner sei Des $= \{b, 1\}$. Wir zeigen für beliebige Formeln A, daß

$$(*) \qquad \forall I \in \mathbf{2} : I(A) = 1 \text{ in } \mathbf{KL} \;\Rightarrow\; \forall I \in \mathbf{3} : I(A) \in \text{Des in } \mathbf{LP}.$$

Wir nehmen das Antezedens an und betrachten eine beliebige dreiwertige Bewertung I. Wir zeigen, daß $I(A) \in$ Des. Ausgehend von I definieren wir eine Bewertung I^* so:

$$I^*(A) = 1 \text{ gdw } I(A) \in \text{Des}.$$

Es ist unschwer zu verifizieren, daß $I^* \in \mathbf{2}$. Dazu müssen wir nur zeigen, daß $I^*(\neg A) = 1$ gdw $I^*(A) = 0$, und daß $I^*(A \wedge B)$ gdw $I^*(A) = 1 = I^*(B)$. Aus dem angenommenen Antezedens von $(*)$ folgt dann, daß $I^*(A) = 1$ und also, nach der Definition von I^*, daß $I(A) \in$ Des. ■

Mit Hilfe der Konsequenzoperationen für **LP** und **KL**, können wir diese Beobachtungen auch so formulieren: Obwohl im allgemeinen, d.h. für alle Formelmengen X, *nicht* gilt, daß $\mathrm{Cn}_{\mathbf{LP}}(X) = \mathrm{Cn}_{\mathbf{KL}}(X)$ haben wir für den speziellen Fall $X = \emptyset$, $\mathrm{Cn}_{\mathbf{LP}}(\emptyset) = \mathrm{Cn}_{\mathbf{KL}}(\emptyset)$.

Nach der Beobachtung sind insbesondere auch die folgenden Formeln gültig in **LP**:

$$A \vee \neg A,$$
$$\neg(A \wedge \neg A),$$
$$A \wedge \neg A \to B.$$

Die Gültigkeit der letzten zwei Formeln in einer parakonsistenten Logik mag zunächst befremden. Wie kann $\neg(A \wedge \neg A)$ gültig sein, wenn wir Widersprüchen die Möglichkeit einräumen möchten, wahr zu sein? Nun, wir können diese Möglichkeit einräumen und zugleich sagen, daß Widersprüche auch immer falsch sind. Wenn wir sagen, daß A widersprüchlich, d.h. zugleich wahr und falsch sei, dann können wir damit nur meinen, daß A den Wert $b = \{w, f\}$ hat. Aber dann haben auch $A \wedge \neg A$ und $\neg(A \wedge \neg A)$ den Wert b – und dieser Wert ist designiert, weil auch in ihm Wahrheit enthalten ist.

Die Gültigkeit von $A \wedge \neg A \to B$ ist merkwürdiger. Aber auch dieser Merkwürdigkeit kann man eine Deutung zugunsten von **LP** geben. **LP** gelingt es offenbar sehr erfolgreich, die Stärke der klassischen Logik soweit zu approximieren wie es mit der Eigenschaft der Parakonsistenz vereinbar ist. Diese verlangt ja nur, daß nicht jede Menge, in der ein widersprüchliches Paar A und $\neg A$ enthalten ist, durch den Abschluß unter $\mathrm{Cn}_{\mathbf{LP}}$ trivialisiert wird. Da auch die DeMorgan-Äquivalenzen und der klassische Zusammenhang zwischen Implikation und Disjunktion in **LP** erhalten bleiben, so findet sich in der Folge auch $A \wedge \neg A \to B$ unter den gültigen Formeln. Das kompromittiert weder die Eigenschaft der Parakonsistenz im engeren Sinne, noch die stärkere Eigenschaft, daß **LP** nicht endlich trivialisierbar sein möge: Für *keine* endliche Menge X ist $\mathrm{Cn}_{\mathbf{LP}}(X)$ die Menge aller Formeln.

Einige klassische Folgerungen sind in **LP** jedoch nicht mehr gültig. Das muß zwangsläufig dazu führen, daß manches doch nicht mehr so funktioniert, wie es uns aus der klassischen Logik vertraut ist. Einiges davon hat sicher damit zu tun, daß **LP**, im Gegensatz zur klassischen Logik, Widersprüche ernstnimmt. Aber gibt es auch Einbußen an logischen Zusammenhängen, welche so nicht genügend motiviert werden können? Muß **LP** logische Kollateralschäden hinnehmen, die auch aus parakonsistenter Sicht eigentlich bedauerlich sind?

Die zunächst erstaunlichste Beobachtung ist, daß Modus Ponens kein gültiger Schluß in **LP** ist. Die zur Folgerungsregel äquivalente Sequenz

MP. $$A, A \to B \vdash B$$

ist in **LP** nicht gültig. Um MP zu widerlegen, betrachte man eine Interpretation I mit $I(A) = b$ und $I(B) = 0$. Dann ist $I(A \to B) = b$. In diesem Fall haben die beiden Prämissen A und $A \to B$ designierte Werte, während die Konklusion B den Wert 0 hat.

Daß Modus Ponens für den Pfeil ("wenn ... dann ... "!) nicht gelten soll, scheint ein nicht hinnehmbarer Eingriff in die Logik zu sein. Aber wir erinnern uns, daß $A \to B$ nach den Tafeln für **LP** nichts anderes bedeutet als $\neg A \vee B$. Modus Ponens ist also nur eine Notationsvariante des disjunktiven Syllogismus. Dieser Schlußmodus ist aber mit der Eigenschaft der Parakonsistenz nicht vereinbar, wie die erste Ableitung (p. 293) gezeigt hat.

Im Abschnitt über die **C**-Systeme haben wir schon einmal dargelegt, warum der disjunktive Syllogismus bei drohenden Widersprüchen keine gute Idee sein kann. Diese Überlegungen können wir auf das dreiwertige Szenario, in dem wir uns nun befinden, gut übertragen. Dazu betrachten wir wieder den DS in der Sequenzen-Form

DS. $$\neg A, A \vee B \vdash B.$$

Angenommen, A ist wahr und falsch. Dann ist auch $\neg A$ wahr. Ebenso ist in diesem Fall $A \vee B$ wahr, gleichgültig, welchen Wert B hat. Insbesondere könnte B kein bißchen wahr, d.h. rein falsch sein. Wenn wir in dieser Situation nach DS schließen, dann schließen wir von (teilweise) wahren Prämissen auf eine völlig falsche Konklusion. Nicht gut. – Im Umfeld drohender Widersprüche kann DS kein gültiger Schluß sein. Wird die Implikation im disjunktiven Sinne aufgefaßt, dann ist MP äquivalent zum DS und darum kann dann auch MP, wie wir sahen, kein gültiger Schluß sein.

Eine weitere Sequenz, die in **LP** nicht gilt drückt die Transitivität der Implikation aus:

TRANS. $$A \to B, B \to C \vdash A \to C.$$

Es sei $I(A) = 1$ und $I(C) = 0$. Dann ist $I(A \to C) = 0$. Wenn jetzt $I(B) = b$, dann nehmen die beiden Prämissen den designierten Wert b an.

Die Ungültigkeit der Transitivitätssequenz läßt sich auf die Ungültigkeit des disjunktiven Syllogismus zurückführen. Denn TRANS ist äquivalent zu der folgenden Sequenz:

$$\neg A \vee B, \neg B \vee C \vdash \neg A \vee C.$$

Nun ist leicht zu sehen, daß das Schließen hier auf dem disjunktiven Syllogismus beruht. Wir beginnen mit der ersten Prämisse, $\neg A \vee B$. Im Falle $\neg A$ können wir gleich auf $\neg A \vee C$ schließen. Im Falle B schließen wir aus der zweiten Prämisse nach DS, daß C und also $\neg A \vee C$. Da es gute Gründe gibt, im parakonsistenten Kontext den disjunktiven Syllogismus abzulehnen, ist die Sequenz TRANS ebenso wie MP abzulehnen – unter der Voraussetzung, daß Implikation und Disjunkion auf klassische Weise verknüpft sind.

Die Beobachtungen über Modus Ponens und Transitivität, die wir gerade gemacht haben, können wir aber auch gegen genau diese Voraussetzung wenden. Genauer gesagt, liegt jetzt der folgende Einwand nahe.

Wenn in einer Logik wie **LP** der disjunktive Syllogismus nicht gilt, dann kann ein Operator $\circ$, definiert durch

(Def) $$A \circ B \; := \; \neg A \vee B$$

nicht die Modus Ponens-Bedingung

($\circ$) aus A und $A \circ B$ schließe auf B

erfüllen. Die Bedingung ($\circ$) ist aber notwendig dafür, daß es sich bei dem Operator $\circ$ um eine Implikation handelt – jedenfalls, wenn wir unter einem Implikationssatz irgendeine Art von indikativischem Wenn-Dann-Satz verstehen. “Wenn A, dann B” sagt so viel wie: “Gegeben, daß A wahr ist, so ist auch B wahr”. Wenn also A gegeben ist, dann sollten wir aus dem Konditional auf B schließen dürfen. Das ist Modus Ponens. Stützt ein zweistelliger Junktor einen solchen Schluß nicht, dann kann er keine Implikation darstellen. Also kann das, was wir bisher in unserer Darstellung von **LP** mit dem Pfeil $\rightarrow$ notiert haben, gar keine Implikation im normalen Sinne sein. In **LP**, definiert durch die Tafeln für $\neg$, $\vee$ und $\wedge$, fehlt eine richtige Implikationsoperation!

Das ist offenbar ein allgemeines Problem für jede Art parakonsistenter Logik, die den disjunktiven Syllogismus ablehnt. Denn in diesem Fall wird keine echte Implikation – d.h. eine, die den Schluß nach MP stützt – einfach per (Def) durch die Bedeutungsfestlegungen für $\neg$ und $\vee$ zur Verfügung gestellt. Vielmehr muß eine solche Implikation separat eingeführt werden. Im folgenden stellen wir zwei Erweiterungen von **LP** um eine echte Implikation vor.

Die Logiken LP$^{\rightarrow}$ und RM3. Um Modus Ponens und auch Transitivität wieder in ihre Rechte einzusetzen, gibt es eine einfache Reparatur. In der Kleene-Tafel für die Implikation (p. 317) müssen wir $b \rightarrow 0$ statt auf b, auf 0 setzen. Damit werden die oben angeführten Widerlegungen von MP und

TRANS ausgeschlossen. Mit diesem, für den Zweck minimalsten Eingriff, entsteht die Logik $\mathbf{LP}^{\rightarrow}$.[21]

Implikation in $\mathbf{LP}^{\rightarrow}$

$\rightarrow$	1	b	0
1	1	b	0
b	1	b	0
0	1	1	1

Die Tafel können wir auch kurz so beschreiben:

$$I(A \rightarrow B) = \begin{cases} 1, & \text{wenn } I(A) = 0; \\ I(B), & \text{anderenfalls.} \end{cases}$$

Oder so:

$$(\mathbf{LP}^{\rightarrow} \rightarrow) \qquad I(A \rightarrow B) \begin{cases} = 0, & \text{falls } A \in \text{Des und } B \notin \text{Des}; \\ \in \text{Des}, & \text{anderenfalls.} \end{cases}$$

Eine Implikation ist also falsch, wenn das Antezedens "gut" und das Konsequens "schlecht" ist; anderenfalls ist die Implikation "gut" – die natürliche Fortschreibung, also, der klassischen Tafel in das dreiwertige Szenario mit den zwei designierten Werten. Durch die Änderung $b \rightarrow 0 = 0$ koinzidieren nun – anders als in $\mathbf{LP}$ – die Theoreme von $\mathbf{LP}^{\rightarrow}$ nicht mehr mit denen der klassischen Logik. Insbesondere ist

$$\text{EFQ.} \qquad A \wedge \neg A \rightarrow B$$

kein Theorem von $\mathbf{LP}^{\rightarrow}$. (Sei $A = b$ und $B = 0$. Dann ist $A \wedge \neg A \rightarrow B = b \rightarrow 0 = 0$.) Auch die Form

$$A \rightarrow (\neg A \rightarrow B)$$

wird ebenso falsifiziert.

[21] Avron in [18] behandelt diese und weitere Variationen über das Thema der Implikation in einer dreiwertigen Logik. Da Logiken wie $\mathbf{LP}^{\rightarrow}$ und **RM3** sehr naheliegende Kandidaten im Reich der dreiwertigen Logik sind, wurden sie von verschiedenen Autoren und oft unabhängig voneinander unter verschiedenen Namen vorgestellt. In Batens [21] erscheint $\mathbf{LP}^{\rightarrow}$ unter dem Namen **PI**. In Avron [18] finden wir **LP** als *Pac*, und $\mathbf{LP}^{\rightarrow}$ heißt dort "*Pac* with internal implication".

Andererseits bewahrt $\mathbf{LP}^{\to}$ eine wichtige Eigenschaft, die in der Logik RM3, die wir gleich besprechen werden, nicht mehr gilt. In $\mathbf{LP}^{\to}$ verhalten sich Implikation und Ableitbarkeit harmonisch zueinander im Sinne des *Deduktionstheorems*:

DED. $$X, A \vdash B \text{ gdw } X \vdash A \to B.$$

Von rechts nach links ist das einfach die Modus Ponens-Regel. Von links nach rechts gelesen, dürfen wir von der Behauptung einer gültigen Ableitung zur Behauptung eines gültigen Konditionals übergehen. Diese Richtung ist einfach zu beweisen.

BEOBACHTUNG 9. *In* $\mathbf{LP}^{\to}$ *gilt: Wenn* $X, A \vdash B$*, dann* $X \vdash A \to B$*.*

BEWEIS. Inspektion des Beweises des Deduktionstheorems für die klassische Aussagenlogik zeigt, daß die Gültigkeit der folgenden Formeln und Ableitungsregeln hinreichend ist:

1. $A \to (B \to A)$;
2. $A \to (B \to C) \to ((A \to B) \to (A \to C))$;
3. Modus Ponens (als alleinige Ableitungsregel!).

Es ist leicht nachzuprüfen, daß $\mathbf{LP}^{\to}$ die Bedingungen 1-4 erfüllt.

Wir können aufgrund der Tafel für $\to$ auch wie folgt argumentieren. Angenommen, $X, A \vdash B$, d.h.

$(*)$ $$\forall I : I(X, A) \in \text{Des} \Rightarrow I(B) \in \text{Des},$$

sowie

$(**)$ $$I(X) \in \text{Des}.$$

(Zu zeigen: $I(A \to B) \in \text{Des}$.) Es folgt aus $(**)$, daß

$(\dagger)$ $$I(A) \in \text{Des} \Rightarrow I(X, A) \in \text{Des}.$$

Nun ist entweder $I(A) \in \text{Des}$ oder $I(A) = 0$. Im ersten Fall folgt aus $(\dagger)$ und $(*)$, daß $I(B) \in \text{Des}$; in welchem Fall die Tafel $I(A \to B) \in \text{Des}$ ergibt. Im zweiten Fall bestimmt die Tafel, daß $I(A \to B) = 1$. ∎

Die Implikationstafel für die Logik **RM3** (siehe [10]) übernimmt die Korrektur für $b \to 0$ von $\mathbf{LP}^{\to}$ und weicht in einem weiteren Punkte von der – durch die Definition $A \to B := \neg A \vee B$ festgelegten – Kleene-Tafel für $\to$ ab: $1 \to b = 0$ (statt b).[22]

[22] Axiomatisierungen von $\mathbf{LP}^{\to}$ und **RM3** findet der Leser z.B. in [18].

Implikation in **RM3**

$\to$	1	b	0
1	1	0	0
b	1	b	0
0	1	1	1

Wie $\mathbf{LP}^{\to}$ übernimmt **RM3** für $\neg$, $\vee$ und $\wedge$ die starken Kleene-Tafeln. Diese Information können wir also weiter im Hasse-Diagramm für die starken Kleene-Tafeln (p. 317) mit der Anordnung

$$(<) \qquad 0 < b < 1$$

darstellen. Aus der **RM3**-Tafel können wir ablesen, daß

$$(\mathbf{RM3} \to) \qquad I(A \to B) \begin{cases} \in \text{Des, wenn } I(A) \leq I(B); \\ = 0, \ \text{anderenfalls.} \end{cases}$$

Das ist eine sehr natürliche Art und Weise, die Implikation in der Werteanordnung $(<)$ zu behandeln. Wenn wir uns noch einmal die Werte 0, b und 1 als $\{f\}$, $\{w, f\}$, und $\{w\}$ vorstellen, so besagt (**RM3** $\to$), daß eine Implikation falsch (0) ist, wenn das Konsequens mehr Falschheit bzw. weniger Wahrheit enthält als das Antezedens; und eine Implikation enthält Wahrheit (hat einen designierten Wert), wenn das Konsequens mindestens so viel Wahrheit enthält wie das Antezedens.[23] In $\mathbf{LP}^{\to}$dagegen, kann $I(A \to B)$ designiert sein, obgleich $I(B) < I(A)$. Das geschieht genau dann, wenn $I(A) = 1$ und $I(B) = b$.

Gegenüber $\mathbf{LP}^{\to}$ erhöht die weitere Festsetzung $1 \to b = 0$ (statt b) die Chance einer Implikation, keinen designierten Wert zu erhalten. Also ist **RM3** ein Teilsystem von $\mathbf{LP}^{\to}$. Modus Ponens und Transitivität sind weiter gültig in **RM3**, denn sie wurden ja durch $b \to 0 = 0$ wieder in ihre Rechte gesetzt. Dagegen ist nun

$$\text{VEQ.} \qquad A \to B \vee \neg B$$

ebensowenig gültig wie die rein implikative Form

$$B \to (A \to B).$$

[23] Die einzigen Alternativen zu **RM3**, die (erste Klausel der) Bedingung ($\to$) zu realisieren, wären $b \to 1 = b$ und $b \to b = 1$.

(Es sei $A = 1$ und $B = b$. Dann $B \vee \neg B = b$ und $A \to B \vee \neg B = \underline{1 \to b = 0}$; bzw. und $A \to B = \underline{1 \to b = 0}$ und also $B \to 0 = b \to 0 = 0$.)

Die Junktoren, die durch die oben angegebenen Tafeln für $\to$ bestimmt sind, können wir die *kanonischen* Implikationen in $\mathbf{LP}^{\to}$ bzw $\mathbf{RM3}$ nennen. In beiden Logiken lassen sich natürlich weitere Junktoren definieren und einige davon sind den jeweiligen kanonischen Implikationen sehr ähnlich. Es stellt sich heraus, daß wir in beiden Logiken die kanonische Implikation der jeweils anderen Logik ausdrücken können. In $\mathbf{LP}^{\to}$ können wir einen (stärkeren) Pfeil $\to^{+}$ so definieren:

$$A \to^{+} B \ := \ (A \to B) \wedge (\neg B \to \neg A).$$

Wenn wir uns die Mühe machen, die Wahrheitstafel für $\to^{+}$ in $\mathbf{LP}^{\to}$ zu errechnen, dann stellen wir fest, daß es genau die Tafel für $\to$ in $\mathbf{RM3}$ ist. In $\mathbf{LP}^{\to}$ können wir also die $\mathbf{RM3}$-Implikation ausdrücken, d.h. $\mathbf{RM3}$ ist unter der Übersetzung in $\mathbf{LP}^{\to}$ enthalten. Umgekehrt gilt das genauso. Denn wenn wir in $\mathbf{RM3}$ einen (schwächeren) Pfeil so defieren:

$$A \to^{-} B \ := \ (A \to B) \vee B,$$

dann ist die Wahrheitstafel für $\to^{-}$ genau die Tafel für $\to$ in $\mathbf{LP}^{\to}$. Wenn wir nur das Fragment $(\mathbf{LP}^{\to})^{+}$ von $\mathbf{LP}^{\to}$ in $\{\neg, \vee, \wedge, \to^{+}\}$ betrachten, dann ist $(\mathbf{LP}^{\to})^{+} = \mathbf{RM3}$; und ebenso ist $\mathbf{RM3}^{-}$ in $\{\neg, \vee, \wedge, \to^{-}\}$ identisch mit $\mathbf{LP}^{\to}$. Die beiden Systeme unterscheiden sich also nur in der jeweils anderen Auszeichnung einer kanonischen Implikation.

Die Schemata EFQ und VEQ (in beiden Formen) sind paradigmatische Beispiele für die sogenannten Paradoxien der materialen Implikation. EFQ drückt etwas aus, was viele parakonsistente Logiker ablehnen, auch wenn es parakonsistente Logiken gibt, die EFQ als Theorem aufweisen und deren Folgerungsrelation inkonsistente Annahmen dennoch nicht trivialisiert – **LP** ist eine solche Logik. Wenn wir *beide* Paradoxien der materialen Implikation ablehnen, wie das in **RM3** der Fall ist, so verlassen wir den Bereich der parakonsistenen Logiken im engeren Sinne und nähern uns dem der Relevanzlogik. Relevanzlogiker lehnen EFQ und VEQ ab, weil hier konditionale Verbindungen ausgedrückt werden in Fällen, in denen das Antezedens ersichtlich nichts mit dem Konsequens zu tun haben muß. Für Relevanzlogiker ist Parakonsistenz ein willkommener Effekt, nicht aber das treibende Motiv. Dieses Motiv müssen sich parakonsistente Logiker offensichtlich nicht zu eigen machen. Aus rein parakonsistenter Sicht geht **RM3** daher mit der Ablehnung von VEQ einen Schritt zu weit. Ohne Not schwächen wir die Logik und verstoßen so gegen da Costas vierte Bedingung.

Ein weiterer Grund, diesen Schritt nicht zu tun, ist das Deduktionstheorem DED. Sobald VEQ nicht mehr zur Verfügung steht, wird es mit dem Beweis des Theorems schwierig. (Im nächsten Kapitel werden wir darauf detailliert eingehen.) Tatsächlich können wir DED in **RM3** leicht widerlegen. Wir wissen, daß $B \vee \neg B$ unter jeder Interpretation einen designierten Wert hat. Also gilt

$$\forall I : I(A) \in \text{Des} \Rightarrow I(B \vee \neg B) \in \text{Des},$$

d.h. $A \vdash B \vee \neg B$. Jedoch nicht $\forall I : I(A \rightarrow B \vee \neg B) \in \text{Des}$. Obgleich sich **RM3** nicht als Relevanzlogik im eigentlichen Sinne erweist (sie verletzt die Eigenschaft der Variablenüberschneidung), zeigt sich hier eine Schwierigkeit unter der auch die Relevanzlogiken leiden und die wir kurz so beschreiben können: Die semantische Folgerungsrelation ist wie die Ableitungsrelation monoton, die Implikation ist es in diesen Systemen nicht.

6. Parakonsistente Logik und andere Black Boxes

In diesem Kapitel haben wir nur die drei wichtigsten Ansätze zur parakonsistenten Logik im engeren Sinne und deren prominenteste Realisierungen vorgestellt. Weitere Ansätze und Variationen findet der Leser in Marcos [208], sowie in [49], [48], [223], und [226]. Solche parakonsistente Logiken, die zugleich Relevanzlogiken sind, werden im nächsten Kapitel behandelt.

Alle parakonsistenten Logiken haben gemein, daß sie eine Operation $\blacksquare$ erzeugen, die eine Input-Menge von Formeln in eine Output-Menge so abbildet, daß $\blacksquare$ endliche inkonsistente Mengen nicht trivialisiert. Damit ist gemeint, daß die Operation so sein sollte, daß es Mengen X gibt mit der Eigenschaft $\{A, \neg A\} \subseteq \blacksquare(X)$ und $\blacksquare(X) \neq \text{FML}$. Besser noch, sollte $\blacksquare$ gar keine endlichen Mengen trivialisieren, d.h.

$\blacksquare$ Tr. wenn X endlich ist, dann $\blacksquare(X) \neq \text{FML}$.

(Vgl. dazu die Ausführungen über endliche Trivialisierbarkeit im Abschnitt über die **C**-Systeme.) In allen hier behandelten parakonsistenten Logiken wird vorausgesetzt, daß in der *Black Box* eine deduktive Folgerungsrelation $\vdash$ auf die einfachste Weise am Werke ist: $A \in \blacksquare X$ gdw $X \vdash A$. Daraus folgt sogleich, daß $\blacksquare$ diese Eigenschaften haben muß:

$\blacksquare$I. $X \subseteq \blacksquare(X)$,

$\blacksquare$II. $\blacksquare(\blacksquare(X)) \subseteq \blacksquare(X)$,

$\blacksquare$III. $X \subseteq Y \Rightarrow \blacksquare(X) \subseteq \blacksquare(Y)$.

Unter diesen Bedingungen kann die Agenda für eine parakonsistente Logik nur darin bestehen, eine geeignete sub-klassische Ableitungsrelation $\vdash$ zu finden.

Anders sieht die Sache aus, wenn wir bereit sind, in der *Black Box* einen komplexeren Mechanismus zu finden. Der Mechanismus könnte auf der klassischen Relation $\vdash$ basieren und deren trivialisierenden Effekt auf irgendeine Weise neutralisieren. Einfache Beispiele eines solchen Mechanismus sind die Schlußoperationen von Rescher und Manor [245]. Zu einer Menge X betrachten wir die Familie $mk(X)$ der maximal konsistenten Teilmengen von X. (Im Fall $X = \{A \wedge \neg A\}$ ist $mk(X) = \{\emptyset\}$.) Dann können wir eine erste Operation so definieren:

$$\blacksquare_1(X) = \mathrm{Cn}(\bigcap mk(X)),$$

wobei Cn die klassische Konsequenzoperation ist. Unter den weiteren Möglichkeiten sind

$$\blacksquare_2(X) = \bigcap\{\mathrm{Cn}(Y) : Y \in mk(X))\}, \text{ sowie}$$
$$\blacksquare_3(X) = \bigcup\{\mathrm{Cn}(Y) : Y \in mk(X))\}.$$

Ohne diese Operationen näher zu untersuchen, seien sie nur kurz kommentiert.

$\blacksquare_1$ nimmt die (klassischen) Konsequenzen solcher Formeln, die allen mk-Mengen gemeinsam sind. (In [245] "freie Konsequenz" genannt.) Die Operation erfüllt offensichtlich $\blacksquare$Tr.

$\blacksquare_2$ generiert zunächst die Konsequenzen jeder mk-Menge und nimmt dann deren Schnitt ("starke Konsequenz"). Der Schnitt einer unter Cn abgeschlossenen Menge ist selbst unter Cn abgeschlossen. Auch diese Operation erfüllt $\blacksquare$Tr.

$\blacksquare_3$ schließt auch erst jede mk-Menge unter Cn ab und erzeugt dann deren Vereinigung ("schwache Konsequenz"). Diese Operation garantiert nicht den Abschluß unter Cn, hat aber ebenfalls die Eigenschaft $\blacksquare$Tr. Das mag zunächst überraschen, aber ein einfaches Beispiel illustriert den Punkt: $mk\{A, \neg A\} = \{\{A\}, \{\neg A\}\}$. In $\mathrm{Cn}(A) \cup \mathrm{Cn}(\neg)$ haben wir dann zwar sowohl A als auch $\neg A$, aber nicht deren Konjunktion und was daraus folgt.

Die dritte Operation ist offenbar eng verwandt mit den schon besprochenen nicht-adjunktiven Logiken Jaśkowskis. Viele weitere Variationen über dieses Thema sind denkbar. Einige davon werden ebenfalls in [245] vorgestellt, andere in dem späteren Buch [244]. Diderik Batens' *Adaptive Logics* [22] geben ähnliche Beispiele ab, wie die *Black Box* mit parakonsistentem Effekt auf der Grundlage klassischer Folgerung gefüllt werden kann; in [23] wird die Beziehung zu den Arbeiten von Jaśkowsi und denen von Rescher und Manor dargestellt. Es ist nur ein kleiner weiterer Schritt, in die *Black Box* eine Präferenzrelation einzufügen, welche die Auswahl konsistenter Teilmengen von Annahmen steuern kann. Damit kommen wir den modernen Theorien anfechtbaren Schließens (siehe Kapitel VII) sehr nahe.

Wie immer die Details aussehen mögen, diese Operationen illustrieren auf einfache Weise einen ganz anderen Ansatz zum Problem, wie aus inkonsistenten Prämissen nicht-trivial geschlossen werden kann. Die direkten Ansätze, die wir im Hauptteil dieses Kapitels vorgestellt haben, schlagen eine Abschwächung der klassischen Logik vor. Die Vorschläge nehmen zumeist die Gestalt eines schwächeren axiomatischen Systems oder einer Erweiterung der Klasse der Modelle, über welche die Gültigkeit von Schlüssen definiert wird, an. Eingriffe in Axiome und Regeln sind die traditionellen Verfahren der Logik, weshalb die direkten Ansätze in "traditionelle nicht-klassische Logiken" münden. Die indirekten Ansätze basieren dagegen auf der klassischen Folgerungsrelation, kombinieren diese aber mit Methoden der Annahmenaufbereitung, die sicherstellen, daß immer nur gefahrlos, d.h. aus konsistenten Prämissen geschlossen wird. In der *Black Box* dieser Ansätze wirkt klassische Konsequenz Cn mit einem Aufbereitungsmodul typischerweise so zusammen, daß ■Tr, nicht aber alle Eigenschaften ■I–III erfüllt sind. So hat zum Beispiel keine der drei Rescher-Manor-Operationen die Inklusionseigenschaft ■I. Der traditionelle Folgerungsbegriff wird also erweitert, weshalb die indirekten Ansätze nicht-traditionelle aber im Kern dennoch klassische Logiken produzieren.

Im Kapitel VII über anfechtbares Schließen werden wir ausschließlich Ansätze kennenlernen, die auf genau diese Weise verfahren, indem sie klassische Folgerung mit Änderungsoperationen über Annahmenmengen kombinieren. Auch dort werden wir für die so definierten Folgerungsoperationen auf interessante Einschränkungen der Bedingungen ■I–III treffen. Parakonsistenz ist jedoch kein unmittelbares Anliegen anfechtbaren Schießens. Dieses ist wesentlich ampliativ (supra-klassisch) und nicht, wie für Parakonsistenz erforderlich, beschneidend (sub-klassisch). Am Ende des Kapitels VIII über *Belief Revision* führen wir jedoch vor, wie sich die Grundideen für den Zweck der Parakonsistenz adaptieren lassen.

VI.

RELEVANZLOGIK

1. Relevanz

Auf der Suche nach einer bisher noch nicht vertretenen Ansicht in der Philosophie des Geistes ist Professor Peters auf die These T gestoßen. [1] Leider ist T schon auf den ersten Blick so abwegig, daß es schwierig sein dürfte, jemanden dazu zu bringen, sich damit zu beschäftigen. Dieser Gefahr tritt Peters so entgegen:

> Entweder ist die These widersprüchlich oder sie ist es nicht. Im ersten Fall folgt aus T, daß es keine naturalistische Reduktion ethischer Urteile geben kann. Im anderen Fall impliziert T, daß ein Gegenstand in der Relation der Identität nur zu sich selbst stehen kann. Gleichgültig also, wie die Diskussion um T entschieden werden sollte, so ist eine Diskussion darüber nicht nur für die Philosophie des Geistes von größtem Interesse.

Einer der Gutachter antwortet darauf, daß er gerne zugestehe, daß in der Philosophie *irgendwie* alles mit allem zusammenhänge, daß er sich aber doch eine etwas detailliertere Auskunft darüber wünsche, wie die sehr allgemeine metaphysische These von der sehr speziellen These in der Philosophie des Geistes abhängig sein könne und wie er sich genau die Beziehung zwischen dieser und der Möglichkeit naturalistischer Reduktion in der Ethik vorstelle. Das Gutachten ist den Gepflogenheiten entsprechend anonym und so wendet sich Peters an die Herausgeberin der Zeitschrift:

[1] Die folgende Geschichte variiert einen Abschnitt aus [10, §2].

Sehr geehrte Kollegin,

ich muß mich sehr über Ihren Gutachter wundern. Meine Studenten lernen bereits im ersten Semester soviel elementare Logik, daß sie sich die Fragen Ihres Gutachters selbst beantworten könnten. Es genügt darauf hinzuweisen, daß, falls die These widersprüchlich ist, sie zwingend zur Ablehnung der genannten Reduktionsmöglichkeit verpflichtet. Ferner können distinkte Gegenstände sicher nicht identisch sein, was sicher auch aus T folgt. [...]

Übung. Antworten Sie Herrn Professor Peters!

Tatsächlich ist es eine Eigenheit der klassischen Logik, daß sie die Konditionale

(efq) Wenn $A \wedge \neg A$, dann B
(eqv) Wenn A, dann $B \vee \neg B$

als gültig ausweist. Sie tut das gleich in zweierlei Sinn: sowohl im Sinne der Gültigkeit der *Formeln*

EFQ. $A \wedge \neg A \to B$

EQV. $A \to B \vee \neg B$,

als auch im Sinne der Gültigkeit der *Folgerungsbehauptungen*

EFsQ. $A \wedge \neg A \vdash B$

EQsV. $A \vdash B \vee \neg B$.

So hat Peters also nicht unrecht, wenn er darauf hinweist, daß seine Bemerkungen im Einklang mit den Gesetzen der elementaren (damit meint er: der klassischen) Logik stehen. Dennoch wirken seine Anwendungen von (efq) und (eqv) auffällig deplaziert. Verteidiger der klassischen Logik weisen darauf hin, daß man auch Wahres deplaziert vorbringen kann. Der Umstand, daß Peters vernünftigerweise gebeten werden sollte, seine Anwendungen von (efq) und (eqv) zurückzunehmen, zeigt also nicht, daß er etwas Falsches behauptet hat.

Diese Art, die klassische Logik zu verteidigen, haben wir in Kapitel 4.6 ("Indikativische Konditionale") ausführlich behandelt. Sie beruht im wesentlichen darauf, daß sich aus der Antwort auf die Frage nach der richtigen Bedeutung eines Satzes nicht ohne weiteres die nach seinem richtigen Gebrauch ergibt. Was hier weiter benötigt wird, sind Gebrauchsregeln, die sich aus allgemeinen Gesprächsmaximen und letztlich aus dem Zweck sprachlichen Austausches ergeben. Die hier skizzierte Verteidigung von (efq) und (eqv) kombiniert eine sehr einfache Theorie der Bedeutung von "Folgern" und "Implizieren" mit einer komplexen Theorie des Gebrauchs solcher Ausdrücke. Aber müssen wir die Linie zwischen Semantik und Pragmatik genau so ziehen? Warum nicht sagen, daß die Bedeutung von "Folgern" und "Implizieren" komplexer ist, als es die klassische Logik sich vorstellt, und daß ein guter Teil der pragmatischen Epitheorie des klassischen Logikers tatsächlich in die Semantik gehört?

Wir beanstanden Peters' Behauptungen , weil sie ganz offensichtlich einen Kompetenzmangel im Hinblick auf "Wenn ..., dann ..." verraten. Aber was bedeutet das genau? Liegt hier ein semantisches oder ein pragmatisches Defizit vor? Hat Peters nicht verstanden, was "Wenn ..., dann ..." bedeutet? Oder ist er ein Gesprächs-Zombie, der nicht weiß, wann was gesagt werden kann? Wenn wir zur Diagnose semantischer Inkompetenz neigen, dann werden wir die Bedeutung von "Wenn ..., dann ..." anders angeben müssen, als es der klassische Logiker tut. Das hat einerseits den Vorteil, daß wir keine pragmatische Epitheorie benötigen, um zu erklären, was mit (efq) und (evq) nicht in Ordnung ist. (Und auf Peters' Brief an den Herausgeber müssen wir uns erst gar nicht einlassen.) Andererseits werden wir uns aber auf eine kompliziertere semantische Theorie einstellen müssen. Es ist genau diese Option, die in der "klassischen" Relevanzlogik (wie wir sie etwa in [10] dargestellt finden) ausbuchstabiert wird.

Schon in den ersten Arbeiten zur Relevanzlogik wird diese Option etwa wie folgt motiviert (siehe z.B. [1] und [12], sehr ausführlich in [10]). Dazu betrachten wir zunächst den Folgerungssinn von (efq) und (eqv), also EFsQ und EQsV. Folgern (oder Ableiten) ist eine Handlung, vielleicht etwas abstrakter als das Zubereiten von Speisen, aber es hat mit diesem gemein, daß es in Schritten vor sich geht und auf ein Ziel zusteuert. Typischerweise setzen wir Annahmen und entwickeln diese schrittweise fort im Hinblick auf eine Konklusion. (Wir stellen die Zutaten bereit und "entwickeln" diese zum fertigen Gericht.) Die Sequenzen EFsQ und EQsV halten jedoch kein Resultat fest, das auf diese, typische Weise erreicht worden ist. Denn in EQsV wurde von der Annahme A gar kein Gebrauch gemacht; das Ziel, $B \vee \neg B$ war gewissermaßen sofort erreicht. (Etwa so, als wenn wir erst frische Zutaten einkaufen und dann doch das fertige Gericht einer Dose entnehmen.) Und in EFsQ wird die Annahme $A \wedge \neg A$ auf gar kein bestimmtes Ziel hin entwickelt; alles dient unterschiedslos und ohne weiteres als Ziel. (Das ist wie eine magische Zutat, aus der jedes Gericht gezaubert werden kann.) Statt zu sagen, daß dieses Folgern "untypisch" sei, sagen Relevanzlogiker, daß es sich dabei um gar kein Folgern handelt. (Wie das Öffnen von

Dosen und Zaubern kein Kochen ist.) Wenn aber EFsQ, und ebenso EQsV, gar kein Resultat von Folgern festhält, dann drücken diese Sequenzen keine gültigen Folgerungsverhältnisse aus. Sie sollten deshalb nicht Bestandteil einer logischen Theorie des Folgerns sein. [2]

Wenn wir die anti-klassische Rhetorik im letzten Absatz einmal ausblenden, dann sehen wir den Umriß einer interessanten Folgerungsbeziehung, die enger als die klassische ist. Wir können es vorerst dahingestellt sein lassen, in welchem Sinne klassische und relevante Folgerung koexistieren können. Wir sollten uns fragen, welche Möglichkeiten es gibt, die Bedingung des echten Gebrauchs von Prämissen präzise zu machen und uns die so entstehende Theorie des relevanten Folgerns ansehen.

Angenommen, wir haben eine relevante Folgerungsbeziehung $\Rightarrow$ fixiert. Dann läßt sich – vielleicht – auf natürliche Weise eine Klasse von Wenn-Dann-Aussagen charakterisieren, für die EFQ und EQV nicht gültig sein können. Die Brücke könnte hier die *Deduktionsäquivalenz*

$$X, A \Rightarrow B \text{ gdw } X \Rightarrow A \to B.$$

bilden. Wenn die Deduktionsäquivalenz gilt, dann spiegelt der Pfeil in der Objektsprache genau den Sinn von “Wenn ..., dann ...” wieder, den wir in der Metasprache durch das Feststellen von (relevanten) Deduktionsverhältnissen ausdrücken: das (relevanz-) logische “Wenn ..., dann ...”. Solche Harmonie zwischen Objekt- und Metasprache scheint erstrebenswert zu sein. Allein, wir werden sehen, daß die Dinge nicht so einfach sind.

Im folgenden werden wir uns einige Versuche ansehen, eine Logik aufzubauen, welche frei ist von solchen Sequenzen und Theoremen, die nicht das Resultat relevanten Folgerns wiedergeben – im oben angedeuteten Sinne. Da in solchen Logiken die Annahmen einer Folgerung (bzw. das Antezedens einer Implikation) “relevant” für die Wahrheit der Konklusion (bzw. des Konsequens) sein sollen, werden diese Logiken *Relevanzlogiken* genannt.

Die Bezeichnung könnte zu der Annahme verführen, daß der Begriff der Relevanz der *Gegenstand* solcher Logiken sei – so etwa, wie der Begriff der Notwendigkeit (in verschiedenen Spielarten) Gegenstand der Modallogik ist. Es ist aber sicher nicht so, daß Relevanzlogiken eine abstrakte, logische Theorie von Relevanz in irgendeinem Sinne anbieten. [3] Relevanz ist nicht

[2] Was wir hier sehr allgemein beschrieben haben, wird in [10] (und anderen Arbeiten) anhand eines Kalküls des natürlichen Schließens vorgeführt. Die Merkwürdigkeit einer Ableitung von EFsQ bzw. von EQsV tritt hier unmittelbar vor Augen. Letztlich ist die sehr ausführliche Betrachtung natürlichen Schließens in [10] aber auch nur eine Intuitionspumpe.

[3] Im Englischen wird Relevanzlogik sowohl als *relevance logic* (meist in Nordamerika),

Gegenstand einer Relevanzlogik, sie deutet vielmehr eine zu erfüllende Bedingung an. Intuitiv ist das die Bedingung, daß alle im Wenn-Teil versammelten Annahmen eines Konditionals für die Wahrheit des Dann-Teils tatsächlich gebraucht werden. Diese Gebrauchsbedingung ist leider ebenso vage wie die zu klärende Beziehung der "Relevanz".

Es gibt jedoch ein scharfes Kriterium, das zumindest terminologisch eine Logik der Familie der Relevanzlogiken zuordnet. Das ist die Bedingung der

> *Variablenüberschneidung*:
> In einer Relevanzlogik ist $A \Rightarrow B$ bzw. $A \rightarrow B$ nur dann eine gültige Sequenz bzw. ein Theorem, wenn es eine Variable gibt, die sowohl in A als auch in B vorkommt.

Anderson und Belnap [10] haben Variablenüberschneidung als eine notwendige Bedingung dafür verstanden, daß im Antezedens und Konsequens eines logischen Wenn-Dann-Verhältnisses nicht von völlig verschiedenen Dingen die Rede ist. Das setzt die nicht unproblematische These voraus, daß wir etwas über die semantische Verwandtschaft zweier Sätze aufgrund eines syntaktischen Vergleichs aussagen können. Ohne diese Voraussetzung haftet der Bedingung der Variablenüberschneidung etwas willkürliches an. Dennoch wird die Bedingung oft als notwendige Bedingung für die Einordnung einer Logik als Relevanzlogik genommen – soweit uns an einer solchen Klassifikation liegt. Die Bedingung schließt sofort aus, daß *ex falso quodlibet* und *ex quodlibet verum* in ihren verschiedenen Erscheinungsformen gelten – was sicherlich eine wesentliche Eigenart von Relevanzlogiken ist. Somit sind alle Relevanzlogiken zugleich parakonsistente Logiken.

Die Bedingung der Variablenüberschneidung sollte nun ihrerseits nicht zu der Vermutung verführen, Relevanzlogik sei das, was übrig bleibt, wenn man die klassisch gültigen Formeln oder Sequenzen durch ein "Relevanzsieb" (wie z.B. die Variablenbedingung) gestrichen hat. Das könnte zwar so sein; d.h. solche Repräsentationsresultate lassen sich nicht ausschließen. [4] Aber sie sind nicht wesentlich für das Unternehmen der Relevanzlogik. Warum einfache Siebversuche zu kurz greifen, werden wir im nächsten Abschnitt sehen. [5]

Die klassische Logik ist eine Theorie, die auf besonders einfache und natürliche Weise leistet, was wir sowohl in semantischer wie in syntaktischer Perspektive von einer Logik erwarten. Eine nichtklassische Logik

als auch als *relevant logic* (in Australien und dem Vereinigten Königreich gebräuchlich) bezeichnet. An letzterer Bezeichnung mag zwar die polemische Spitze stören, aber sie hat zumindest den Vorteil, nicht das im Text genannte Mißverständnis einzuladen.

[4] Die Arbeiten von Makinson [204, 203, 205] sind eine Bemühung in dieser Richtung.

[5] Eine Familie von Logiken, auf die das Siebbild eher zutrifft, ist die der *Relatedness Logics*; siehe [285] and [73].

kann nur dann als Konkurrenzentwurf unser Interesse wecken, wenn sie sich ebenfalls in möglichst einfacher und natürlicher Weise darstellen läßt. [6] Was der Darstellung einer formalen Theorie und damit der Theorie selbst die gewünschte Einfachheit und Natürlichkeit verleiht, läßt sich allgemein schwerlich sagen. Aber wir erkennen diese Eigenschaften, wenn wir Kandidaten betrachten. So ist zum Beispiel aus der Perspektive der relationalen Semantik, das System **K** ein natürlicher Ausgangspunkt und das System **S5** ein natürlicher Endpunkt für die Modallogik. Die im letzten Kapitel behandelten parakonsistenten **C**-Systeme beginnen zwar mit einer bestechend einfachen Idee, haben dann aber Eigenschaften – insbesondere die Substitutionsfehler –, die diese schließlich als grob unnatürlich erscheinen lassen. Ebenfalls bestechend ist die Grundidee von Priest's **LP**, auf die gut motivierten starken Kleene-Tafeln zurückzugreifen und nur die Menge der designierten Werte zu ändern. Aber dann widerspricht die definierte "Implikation" in **LP** unseren berechtigten Erwartungen so sehr, daß wir zu einem einfachen Eingriff in die Tafeln gedrängt werden. So entstehen zwei Logiken, $\mathbf{LP}^{\rightarrow}$ und **RM3**, die als parakonsistente Logiken ob ihrer Einfachheit keine schlechte Figur machen. Dem Leser werden sicher weitere Beispiele dieser Art einfallen.

Wir werden hier vorzugsweise eine semantische Perspektive einnehmen. Wir betrachten zunächst das Folgern in einer Sprache ohne Implikation, d.h. nur mit $\neg$, $\vee$ und $\wedge$, und versuchen daraus die *Quodlibet*-Sequenzen EFQ und EQV auf natürliche Weise zu verbannen. Die resultierende Theorie ist die DeMorgan-Logik (auch bekannt als *First Degree Entailment*). Dann wenden wir uns der reinen Theorie der Implikation zu und versuchen, auf möglichst einfache Weise, implikative *Quodlibet*-Formeln wie

$$A \to (B \to A)$$

auszuschließen. Daran anschließend nehmen wir die Negation so hinzu, daß

$$A \to (\neg A \to B)$$

nicht mehr gültig ist. Schließlich werden wir die Theorien der Implikation und Negation mit der DeMorgan-Logik kombinieren um zum Standardsystem **R** zu gelangen.

Das war die Skizze einer *via negativa*: An jedem Punkte gilt es etwas zu vermeiden. Aber derselbe Weg läßt sich auch positiv darstellen. In

[6] Makinson drückt diese Einstellung pointiert so aus: "In the author's opinion, in the realm of nonclassical logic where nothing has usefulness to recommend it, the only justification for existence is beauty, [...]" [193, p. 41]. Allerdings scheint Makinson hier den Nutzen nichtklassischer Logik zu unterschätzen.

der DeMorgan-Logik wird es uns darum gehen, die grundlegenden logischen Beziehungen zwischen Konjunktion, Disjunktion und Negation festzuhalten (§§2–3). Dann suchen wir eine ebenso unkontroverse Grundlage in der Theorie der Implikation (§§4–5) und, darauf aufbauend, der Negation (§6). Schließlich hoffen wir durch eine natürliche Kombination dieser Theorieelemente zu einer genügend starken Relevanzlogik zu kommen (§§7–8).

In manchen Darstellungen der Relevanzlogik – besonders in einführenden – besteht ein Mißverhältnis zwischen Bestellung und Lieferung. Es kann leicht der Eindruck entstehen, daß die Relevanzlogik eine Theorie "echten" logischen Folgerns verspricht, d.h. eines solchen Folgerns aus Annahmen, bei dem von letzteren auf gehaltvolle Weise Gebrauch gemacht wird. Geliefert wird dann aber (gewöhnlich) eine Theorie logisch gültiger Implikationen und die Relation des logischen Folgerns tritt (meist) in den Hintergrund. Dafür gibt es einen Grund, der Thema in Abschnitt 9 über Deduktionstheoreme in der Relevanzlogik sein wird.

2. DeMorgan-Logik (DML)

Ein naheliegender Versuch. Auch aus klassischer Sicht sind Sequenzen wie EFsQ und EQsV auffällig: Sie sind auf *leere* Weise gültig, d.h. allein aufgrund der Widersprüchlichkeit der Annahme bzw. der logischen Wahrheit der Konklusion. Auch von den Sequenzen

$$A \wedge \neg A \models A \quad \text{und} \quad A \models A \vee \neg A \tag{59}$$

könnte man dies sagen. Diese instanziieren jedoch zugleich Sequenzen, die auf nichtleere Weise gültig sind, nämlich

$$A \wedge B \models A \quad \text{bzw.} \quad A \models A \vee B. \tag{60}$$

Das unterscheidet die "schlechten" *Quodlibet*-Sequenzen EFsQ und EQsV von den "guten" in (1), in denen das Konsequens eben nicht *quodlibet* ist.

Diese Beobachtung legt den Versuch nahe, auf der Basis der klassischen Folgerungsrelation $\models$ eine strengere Folgerungsrelation $\models'$ zu definieren, welche Sequenzen wie (59) von den *Quodlibet*-Sequenzen abtrennt und letztere ausschließt.[7]

$A \models' B$ gdw A und B Instanzen von Formeln A' und B' sind, so daß
$A' \models B'$, und weder
A' kontradiktorisch noch
B' tautologisch ist.

[7] Eine solche Relation wird in Makinson [193, Kap. 2.2] "subtending implication" genannt. In Burgess [43, Kap. 5.3] wird dieselbe Relation unter dem Namen "perfect entailment" behandelt. Tennant [278] enthält einen ähnlichen Vorschlag. Die Verallgemeinerung für den Fall einer *Menge* von Annahmen muß uns hier nicht interessieren.

Die klassische Folgerungsrelation $\models$ wird hier durch das Sieb einer Bedingung für A und B gestrichen. Es ist klar, daß diese Definition den gewünschten Keil zwischen die Quodlibet-Sequenzen und Sequenzen wie (59) treibt. Bei näherer Untersuchung stellt sich jedoch schnell heraus, daß $\models'$ einige sehr merkwürdige Eigenschaften hat und daher nicht die natürliche Relation sein kann, die wir suchen.

1. Die Relation $\models'$ ist nicht transitiv. So haben wir beispielsweise

$$A \wedge \neg A \models' (A \vee B) \wedge \neg A$$

(Instanz von $A \wedge C \models (A \vee B) \wedge C$). Ebenso

$$(A \vee B) \wedge \neg A \models' B$$

(auf nicht-leere Weise gültig). Jedoch nicht: $A \wedge \neg A \models' B$.

2. Die Relation $\models'$ ist abhängig von der Wahl primitiver Junktoren. Aus der Wahrheitstafel für $\leftrightarrow$ geht hervor, daß $A \leftrightarrow B$ genau dann wahr ist, wenn A und B identische Werte haben. Daher ist $A \leftrightarrow A$ tautologisch, weshalb

$$A \leftrightarrow B \models' A \leftrightarrow A \tag{61}$$

offensichtlich nicht richtig sein kann. Wenn wir aber $A \leftrightarrow B$ definitorisch als Abkürzung für $(A \rightarrow B) \wedge (B \rightarrow A)$ auffassen, dann müssen wir die Frage, ob (61) richtig ist, in die Frage übersetzen, ob

$$(A \rightarrow B) \wedge (B \rightarrow A) \models' A \rightarrow A \tag{62}$$

richtig ist. Diese Sequenz (62) ist jedoch richtig, denn sie instanziiert die Sequenz

$$(A \rightarrow B) \wedge (B \rightarrow C) \models A \rightarrow C$$

und erfüllt die anderen Bedingungen der Definition. Mit $\leftrightarrow$ auf die übliche Weise definiert, stellt sich (61) als richtig heraus; fassen wir dagegen $\leftrightarrow$ als primitiven Junktor auf, dann ist (61) falsch.

3. Klassisch gilt:

$$\text{Wenn } A \vee B \models C, \text{ dann } A \models C; \text{ und} \tag{63}$$

$$\text{wenn } A \models B \wedge C, \text{ dann } A \models B. \tag{64}$$

Für die Relation $\models'$ gilt jedoch weder das eine noch das andere. Denn wir haben

$$(A \wedge \neg A) \vee B \models' B, \tag{65}$$

jedoch nicht $A \wedge \neg A \models' B$. Ebenso haben wir

$$A \models' (B \vee \neg B) \wedge A \tag{66}$$

jedoch nicht $A \models' B \vee \neg B$. Problematisch sind hier die Sequenzen (65) und (66). Im ersten Fall entgeht das Antezedens der Bedingung keine Kontradiktion zu sein durch das disjungierte Anhängsel B; im zweiten Fall entgeht das Konsequens der Bedingung keine Kontradiktion zu sein durch das konjungierte Anhängsel A. Neben den trivialen Sequenzen der Form $A \models A$ stehen aber letztlich hinter der Richtigkeit von (65) und (66) die Quodlibet-Sequenzen EFsQ und EQsV. Auch nach der strengeren Definition der Folgerung bleiben diese Sequenzen also auf unerwünschte Weise wirksam.

Die Beispiele zeigen, daß eine Interpretation der Relation $\models'$ im Sinne logischer Folgerung problematisch ist. Von der Transitivität des Folgerns machen wir zum Beispiel in diesem Buch ständig Gebrauch: Wir beweisen erst Lemmata, um sie dann zu verketten und über die Transitivität solcher Ketten auf Sätze zu schließen. Es ist uns auch – aus logischer Perspektive – gleich, welche von einer Auswahl bedeutungsgleicher sprachlicher Formulierungen wir für die Lemmata und Sätze wählen. Folgerung sollte unter Ersetzung bedeutungsgleicher logischer Partikel invariant sein. Schließlich ist kein unabhängiger Grund erkennbar, warum wichtige Folgerungsprinzipien wie (63) und (64) zugunsten merkwürdiger Sequenzen wie (65) und (66) aufgegeben werden sollten. Eine zunächst naheliegende Idee erweist sich so als Sackgasse.

Lücken und Überschneidungen: Der DeMorgan-Diamant. Klassisch gesehen, folgt B aus $A \wedge \neg A$, weil $A \wedge \neg A$ nie wahr sein kann. Ebenso folgt $B \vee \neg B$ aus A, weil $B \vee \neg B$ immer wahr sein muß. Die eine Begründung versagt jedoch, sobald wir die Möglichkeit in Betracht ziehen, daß $A \wedge \neg A$ manchmal wahr sein kann. Und die andere Begründung versagt, wenn wir nicht länger annehmen, daß $B \vee \neg B$ nie unwahr sein kann. Mit anderen Worten, die klassische Logik beruht auf zwei Annahmen, nämlich, daß es zu keinen Wahrheitswertüberschneidungen (A und $\neg A$ beide wahr) und zu keinen Wahrheitswertlücken (weder A noch $\neg A$ wahr) kommen kann. Im letzten Kapitel haben wir dreiwertige Logiken kennengelernt, in denen der "mittlere" Wert jeweils eine Überschneidung bzw. Lücke repräsentiert und so die zwei klassischen Annahmen aufgehoben sind. Wenn wir diese dreiwertigen Logiken zu einer vierwertigen Logik geschickt kombinieren, dann müßten wir dem Ziel nahekommen die *Quodlibet*-Sequenzen auf natürliche Weise auszuschließen.

Die Behandlung der Implikation in den dreiwertigen Logiken, die auf den starken Kleene-Tafeln basiert, bereitet Schwierigkeiten, wie wir im letzten Kapitel gesehen haben. Deshalb werden wir uns zunächst auf das Folgern in Sprachen mit den Junktoren $\neg$, $\wedge$ und $\vee$ konzentrieren. D.h. die Sprachen, die wir jetzt betrachten, verfügen nur über diese drei Junktoren und ggf. solche, die wir daraus definieren möchten; alle Formeln sind "in $\{\neg, \wedge, \vee\}$" wie man auch kurz sagt. Wir wollen unsere Aufgabe jetzt also so auffassen, daß wir eine *metasprachliche Relation* (der Folgerung) untersuchen wollen, deren linke Koordinate aus einer Menge von Formeln in $\{\neg, \wedge, \vee\}$ und deren rechte Koordinate aus einer Formel, ebenfalls in $\{\neg, \wedge, \vee\}$, besteht.

Wir könnten die Sache auch ein wenig anders angehen. Eine Formel $A \rightarrow_1 B$ sei eine *Implikation erster Stufe*, wenn A und B Formeln in $\{\neg, \wedge, \vee\}$ sind. Dann könnten wir nach einer guten Theorie der logischen Wahrheiten in $\{\neg, \wedge, \vee, \rightarrow_1\}$ fragen, wobei logische Wahrheiten der Form $A \rightarrow_1 B$ einfach wiedergeben sollen, daß B aus A logisch folgt. D.h.,

$$(*) \qquad A \rightarrow_1 B \text{ gdw } A \models B;$$

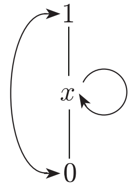
endliche Annahmen*mengen* können wir dabei durch die Konjunktion ihrer Elemente repräsentieren. Diese beiden Ansätze sind ersichtlich äquivalent – ja, im endlichen Fall lediglich Notationsvarianten. Was wir gleich als DeMorgan-Logik vorstellen werden, präsentieren daher andere Autoren als die Logik der erststufigen Implikation (*First Degree Entailment*, FDE). Darin folgen sie einer in [10] begonnenen Tradition. Gewöhnlich gehen wir in der Logik jedoch davon aus, daß eine Satzoperation (ein Junktor) Sätze erzeugt, auf die sie wiederum angewandt werden kann: Satzoperationen können iteriert werden. In diesem Sinne ist $\rightarrow_1$ keine Satzoperation. Auch werden wir später sehen, daß die Vorstellung konsequenter ist, eine Folgerungsrelation in $\{\neg, \wedge, \vee\}$ zu einer in $\{\neg, \wedge, \vee, \rightarrow\}$ zu erweitern als die syntaktische Beschränkung der Implikation auf die erste Stufe aufzuheben. Deshalb ziehen wir hier die relationale Darstellung vor.

Ausgangspunkt unserer Überlegung sind wieder die *starken Kleene-Tafeln* für $\neg$, $\wedge$ und $\vee$ (rechts das charakteristische Hasse-Diagramm):

$\neg$	
1	0
x	x
0	1

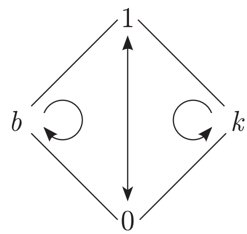
$\wedge$	1	x	0
1	1	x	0
x	x	x	0
0	0	0	0

$\vee$	1	x	0
1	1	1	1
x	1	x	x
0	1	x	0



Wie im letzten Kapitel erklärt, stellt das Hasse-Diagramm die Werte als geordnet vor ($0 < x < 1$). Danach ist $y \wedge z$ immer das Minimum und $y \vee z$ immer das Maximum von (y, z). Die Konjunktion "geht" im Diagramm also nach unten, die Disjunktion nach oben. Die Pfeile deuten die Vorschrift der Negationstafel an.

Wir ziehen jetzt zwei Kopien des Hasse-Diagramms, einmal mit dem Überschneidungswert b ("beides") und einmal mit dem Lückenwert k ("keines") für x. Diese Kopien knicken wir, bildlich gesprochen, an den Mittelwerten ein und fügen sie zum sogenannten *DeMorgan-Diamanten* zusammen:



Der Diamant faßt die folgenden Tafeln zusammen:

$\neg$	
1	0
b	b
k	k
0	1

$\wedge$	1	b	k	0
1	1	b	k	0
b	b	b	0	0
k	k	0	k	0
0	0	0	0	0

$\vee$	1	b	k	0
1	1	1	1	1
b	1	b	1	b
k	1	1	k	k
0	1	b	k	0

Folgerung ist Wahrheitsübertragung von den Prämissen auf die Konklusion. Hier haben wir wieder zwei Möglichkeiten:

(a) Es wird die "reine" Wahrheit, d.h. der Wert 1 übertragen; oder
(b) Es wird Wahrheit im Sinne der Werte 1 oder b übertragen.

Wenn wir die Option (a) wählen, dann wird zwar EQsV ungültig (denn $B \vee \neg B$ kann den undesignierten Wert k annehmen), aber EFsQ bleibt gültig (denn $A \wedge \neg A$ kann den einzigen desiginierten Wert 1 nicht annehmen). Entscheiden wir uns dagegen für (b), d.h. zeichnen wir 1 und b als designierte Wert aus, dann werden beide *Quodlibet*-Sequenzen wie gewünscht ungültig. Denn wenn wir A mit b bewerten, dann erhält $A \wedge \neg A$ ebenfalls den designierten Wert b. Für die *DeMorgan-Logik*, **DML**, definieren wir daher die Folgerungsrelation so (es sei Des die Menge der desgnierten Werte):

$X \models A$ in **DML** genau dann, wenn
für alle Interpretationen I im DeMorgan-Diamenten mit Des $= \{1, b\}$ gilt:
Wenn für alle $A_i \in X$, $[\![A_i]\!]_I \in$ Des, dann $[\![A]\!]_I \in$ Des.

Im nächsten Abschnitt beschränken wir uns auf die Betrachtung von **DML**-Folgerungen aus Einermengen, also im Effekt auf $\models \subseteq$ FML $\times$ FML.

Variablenüberschneidung. In diesem Abschnitt beschränken wir uns auf die Betrachtung von **DML**-Folgerungen aus Einermengen, also im Effekt auf $\models \subseteq$ FML$\times$FML. Durch Prüfung im DeMorgan-Diamanten können wir uns schnell davon überzeugen, daß in **DML** alle DeMorgan-Äquivalenzen erhalten bleiben – daher der Name. (Äquivalenz hier im Sinne von $\equiv$, d.h. $A \equiv B$ gdw $A \models B$ und $B \models A$.) Deshalb haben in **DML** wie in der klassischen Logik alle Formeln äquivalente Normalformen.

Erinnerung. Eine Formel A ist in DNF, wenn A eine Disjunktion von Konjunktionen von Literalen (Atomen oder deren Negationen) ist. Ein Formel A in DNF hat also die Struktur

$$(L_1 \wedge \cdots \wedge L_k) \vee \cdots \vee (L_m \wedge \cdots \wedge L_n),$$

wobei jedes L_i ein Literal, d.h. von der Form P oder $\neg P$ ist. Eine KNF ist, dual, eine Konjunktion von Disjunktionen von Literalen. NB: Im Grenzfall kann eine Disjunktion oder eine Konjunktion aus nur einem Disjunkt bzw. Konjunkt bestehen. Daher ist jedes Literal Grenzfall einer KNF als auch einer DNF.

Im folgenden wollen wir nur solche DNF'en und KNF'en einer Formel betrachten, die auf eine bestimmte Weise aus dieser Formel erzeugt werden können. Dabei soll in einer Formel nur nach folgenden Äquivalenzen (alle gültig in **DML**) ausgetauscht werden ($\bullet$ variabel für $\wedge$ und $\vee$):

Doppelte Negation: $\neg\neg A \equiv A$.
Kommutation: $A \bullet B \equiv B \bullet A$.
Assoziativität: $A \bullet (B \bullet C) \equiv (A \bullet B) \bullet C$.
Distributivität: $A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C)$ und

$$A \vee (B \wedge C) \equiv (A \vee B) \wedge (A \vee C).$$

DeMorgan: $\neg(A \wedge B) \equiv \neg A \vee \neg B$ und $\neg(A \vee B) \equiv \neg A \wedge \neg B$.

Für Normalformen, die ausschließlich nach diesen Regeln erzeugt wurden, verwenden wir die Notation DNF* und KNF*. Eine Formel kann mehr als eine Normalform in diesem Sinne haben. Diese unterscheiden sich jedoch nicht in der Menge der darin vorkommenden Literale (nur in ihrer Anordnung und Assoziierung). Für das gleich zu beschreibende Entscheidungsverfahren, ist es daher gleichgültig welche Normalform* einer Formel wir betrachten.

SATZ 1. *Zu jeder Formel A gibt es eine in* **DML** *äquivalente DNF*, sowie eine äquivalente KNF*.*

BEWEIS. Induktion über den Formelaufbau wie für **KL**. Es wird ausschließlich Gebrauch gemacht von den gerade genannten Umformungen.
∎

Es sei $A \models B$ nun eine in **DML** gültige Sequenz. Nach dem Lemma können wir A und B durch Normalformen A^D (in DNF*) und B^K (in KNF*) ersetzen und erhalten die ebenfalls gültige Sequenz

(67) $$A^D \models B^K.$$

Im einfachsten nichttrivialen Fall ist $A^D = K_1 \vee K_2$ und $B^K = D_1 \wedge D_2$, d.h. (67) sieht so aus:

(68) $$K_1 \vee K_2 \models D_1 \wedge D_2.$$

Diese Sequenz gilt dann und nur dann, wenn

(69) $$\text{(a) } K_1 \models D_1 \wedge D_2 \quad \text{und} \quad \text{(b) } K_2 \models D_1 \wedge D_2.$$

Und (a) gilt dann und nur dann, wenn

(70) $$\text{(aa) } K_1 \models D_1 \text{ und } \text{(ab) } K_1 \models D_2.$$

Gleiches gilt für (69b). Nun ist K_1 eine Konjunktion von Literalen und D_1 eine Disjunktion von Literalen. D.h. (70aa) sieht so aus:

(71) $$L_1 \wedge \cdots \wedge L_m \models M_1 \vee \cdots \vee M_n.$$

Gleiches gilt für (70ab). Für die Gültigkeit von (71) in **DML** – wie auch in **KL** – ist es sicherlich hinreichend, wenn ein L_i $(1 \leq i \leq m)$ identisch ist mit einem M_j $(1 \leq j \leq m)$. Klassisch ist das keine notwendige Bedingung, wie z.B. $P \wedge \neg P \models Q$ sogleich deutlich macht. In **DML** ist die Bedingung im Gegensatz zu **KL** jedoch auch notwendig, wie das folgende Lemma zeigt.

Lemma 2. *Es sei L eine Konjunktion und K eine Disjunktion von Literalen.* Lit(K) *bezeichne die Menge der in L (K) konjungierten (disjungierten) Literale. Wenn $L \models K$ in* **DML**, *dann* $\mathrm{Lit}(L) \cap \mathrm{Lit}(K) \neq \emptyset$.

Beweis. Angenommen $\mathrm{Lit}(L) \cap \mathrm{Lit}(K) = \emptyset$. Dann ist die folgende Bewertung v im DeMorgan-Diamanten möglich: Für alle in L oder K vorkommenden Atome P,

$$I(P) = \begin{cases} 1, \text{ wenn } P \in \mathrm{Lit}(L) \text{ und } \neg P \notin \mathrm{Lit}(L), \\ 0, \text{ wenn } \neg P \in \mathrm{Lit}(L) \text{ und } P \notin \mathrm{Lit}(L), \\ b, \text{ wenn } P, \neg P \in \mathrm{Lit}(L); \\ 0, \text{ wenn } P \in \mathrm{Lit}(K) \text{ und } \neg P \notin \mathrm{Lit}(K), \\ 1, \text{ wenn } \neg P \in \mathrm{Lit}(K) \text{ und } P \notin \mathrm{Lit}(K), \\ k, \text{ wenn } P, \neg P \in \mathrm{Lit}(K). \end{cases}$$

$[\![\ ]\!]_I$ sei die Erweiterung von v nach den Regeln des DeMorgan-Diamanten auf alle Formeln. Eine Induktion über die Längen von L und K zeigt, daß $[\![L]\!]_I$ einer der designierten Werte 1 oder b ist, während $[\![K]\!]_I$ einer der nicht designierten Werte k oder 0 ist. Also ist die Sequenz $L \vdash K$ nicht gültig in **DML**. ∎

Wir können unseren Gedankengang so zusammenfassen:

· Wenn $A \models B$ in **DML**, dann gibt es äquivalente Normalformen* A^D und B^K so, daß $A^D \models B^K$ der Fall ist (und umgekehrt).

· Wenn $A^D \models B^K$, d.h. etwas von der Form

$$C_1 \vee \cdots \vee C_m \models D_1 \wedge \cdots \wedge Dn$$

der Fall ist, dann folgt aus jedem Disjunkt C_i in A^D jedes Konjunkt D_j in B^K (und umgekehrt). (Das verallgemeinert die vereinfachte Betrachtung oben.)

· Alle diese Folgerungen $C_i \models D_j$ sind von der Form (L_i und M_i Literale)

$$L_1 \wedge \cdots \wedge L_m \models M_1 \vee \cdots \vee M_n.$$

· Nach dem Lemma gilt eine solche Folgerung in **DML** nur dann, wenn mindestens ein L_i und ein M_j identisch sind. Die Umkehrung gilt trivialerweise.

Es folgt, daß $A \models B$ in **DML** nur dann gilt, wenn die Bedingung der Variablenüberschneidung (p. 335) erfüllt ist. (Mit etwas mehr Aufwand läßt sich auch die Verallgemeinerung zu $\models\ \subseteq \wp(\mathrm{FML}) \times \mathrm{FML}$ zeigen: $X \models A$ gilt nur dann, wenn X und A Variablen gemein haben.)

Tatsächlich ergeben unsere Überlegungen ein einfaches Entscheidungsverfahren für **DML**. Gegeben eine Sequenz $A \models B$, betrachten wir das Paar (A^D, B^K), wobei A^D eine beliebige DNF* von A und B^K eine beliebige KNF* von B ist. Wenn nun jedes Paar (C, D) von Disjunkten C in A^D und Konjunkten D in B^K so ist, daß die Menge der Literale in C sich nichtleer mit der Menge der Literale in D schneidet, dann ist $A \models B$ gültig in **DML**, anderenfalls nicht.[8]

Einfache Sequenzen wie EFsQ und EQsV fallen sofort durch das Raster dieses Tests. Hier ist eine weitere klassisch gültige Sequenz, die sich als ungültig in **DML** erweist:

> *Disjunktiver Syllogismus*: $P \wedge (\neg P \vee Q) \models Q$. Durch Distribuition erhalten wir die Sequenz in Normalform* $(P \wedge \neg P) \vee (P \wedge Q) \models Q$. Da das erste Disjunkt das Konsequens Q nicht enthält, ist die Sequenz ungültig in **DML**.

Der Disjunktive Syllogismus ist der kritische Schritt im "berüchtigten" Lewis-Argument für EFsQ:

$$\dfrac{\dfrac{\dfrac{A \wedge \neg A \models A \quad A \models A \vee B}{A \wedge \neg A \models A \vee B} \qquad A \wedge \neg A \models \neg A}{A \wedge \neg A \models (A \vee B) \wedge \neg A}\,\mathrm{DS} \qquad (A \vee B) \wedge \neg A \models B}{A \wedge \neg A \models B}$$

Bis auf den Disjunktiven Syllogismus DS sind alle Sequenzen gültig und alle Übergänge (Transitivität von $\models$ und $\wedge$-Einführung) zulässig in **DML**.

Axiomatik. Das folgende Axiomensystem für **DML** (siehe die Tafel unten) entnehmen wir [254, Kap. 2, §§5–6]; vgl. a. [10, §15.2]. Es zeigt einige Schlüsseleigenschaften von DML. Auf einen Nachweis der Richtigkeit und Vollständigkeit, d.h.

$$\Vdash \quad = \quad \models \text{ in } \mathbf{DML}.$$

verzichten wir an dieser Stelle.

[8] Paare (C, D), die diese Bedingung erfüllen heißen (gültige) "einfache Folgerungen" in [10, §15.1]; Paare (A^D, B^K), die sich wie im Test beschrieben auf einfache Folgerungen zurückführen lassen, heißen dort "explizit tautologische Folgerungen". Die gültigen **DML**-Sequenzen sind also genau solche Folgerungen.

Das System **DML**

Annahme	$A \Vdash A$
Schnitt	$\dfrac{X \Vdash A \qquad Y, A \Vdash B}{X, Y \Vdash B}$
Monotonie	$\dfrac{X \Vdash A}{X, Y \Vdash A}$
$\wedge$ Ein	$\dfrac{X \Vdash A \quad X \Vdash B}{X \Vdash A \wedge B}$
$\wedge$ Aus	$A \wedge B \Vdash A \qquad A \wedge B \Vdash B$
$\vee$ Ein	$A \Vdash A \vee B \qquad B \Vdash A \vee B$
$\vee$ Aus	$\dfrac{X, A \Vdash C \quad X, B \Vdash C}{X, A \vee B \Vdash C}$
$\neg\neg$ Ein/Aus	$A \Vdash \neg\neg A \qquad \neg\neg A \Vdash A$
Kontraposition	$\dfrac{A \Vdash B}{\neg B \Vdash \neg A}$

Wenn wir die letzte Regel in der Tafel, Kontraposition, durch die stärkere Regel

Antilogismus. $$\dfrac{X, A \Vdash B}{X, \neg B \Vdash \neg A}$$

ersetzen, dann erhalten wir eine vollständige Axiomatisierung der Konsequenzrelation der klassischen Aussagenlogik **KL**. Hier sind die Herleitungen der *Quodlibet*-Sequenzen in dem durch den Antilogismus verstärkten System:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{B \Vdash B}\,\text{Ann.}}{B, \neg A \Vdash B}\,\text{Monot.}}{B, \neg B \Vdash \neg\neg A}\,\text{Antil.}}{B, \neg B \Vdash A}\,\text{DN} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\cdots}{\neg B, B \Vdash \neg A}}{\neg B \wedge B \Vdash \neg A}\,\text{LinkeAdjunktion}}{A \Vdash \neg(\neg B \wedge B)}\,\text{Kontrap., DN}}{A \Vdash B \vee \neg B}\,\text{DeMorgan, DN}$$

(DN steht für die Streichung doppelter Negationen; Adjunktion der Prämissen folgt aus $\wedge$Aus und Schnitt; die DeMorgan-Äquivalenzen sind etwas umständlicher ableitbar.)

Man könnte der Auffassung sein, daß in dem entscheidenden, linken Beweis nicht der Antilogismus, sondern die Monotonieregel gegen den Geist einer Relevanzlogik verstößt. Wird hier nicht eine für das Folgerungsziel B irrelevante Annahme $\neg A$ hinzugefügt? Die Monotonieregel erlaubt die Ableitung jeder Sequenz der Form $A, B \Vdash A$. In einer solchen Sequenz ist B sicher unnötig für den Schluß auf A.

Dieser Einwand wird uns im Abschnitt über die relevante Implikationstheorie weiter beschäftigen müssen. An dieser Stelle sei schon das grundlegende Problem angedeutet: Wenn $A, B \Vdash A$ eine gültige Sequenz ist, dann würde die zweimalige Anwendung der Deduktionsäquivalenz (p. 334) sofort $A \to (B \to A)$ erzeugen – welches aber relevanzlogisch nicht erwünscht ist; also darf $A, B \Vdash A$ oder die Deduktionsäquivalenz nicht gelten.

Auch wenn die Monotonieregel das Einfügen überflüssiger Prämissen erlaubt, so dürfen wir doch beobachten, daß sie im Einklang mit der Überschneidungsbedingung steht: Erfüllt die Prämissensequenz eines Monotonieschritts die Bedingung, dann tut es trivialerweise auch die Konklusionssequenz. Genau das ist beim Antilogismus nicht der Fall. Denn die Sequenz $X, A \Vdash B$ könnte die Überschneidungsbedingung allein dadurch erfüllen, daß X und B Atome gemein haben. Aber dann würde die Sequenz $X, \neg B \Vdash \neg A$ die Bedingung nicht mehr erfüllen. So ist es tatsächlich nicht die Monotonieregel, sondern der Antilogismus, der durch das Raster der Überschneidungsbedingung fällt. Einwände gegen Monotonie werden also geltend machen müssen, daß die Überschneidungsbedingung nicht hinreichend ist für eine relevante Folgerungsbeziehung.

Die Beobachtung im letzten Absatz findet natürlich ihr Gegenstück im DeMorgan-Diamanten. Wenn unter der Annahme, daß alle Formeln in X designiert sind, auch der Wert von A designiert ist, dann ist der Wert von A designiert unter jeder stärkeren Annahme – also auch unter der Annahme, daß alle Werte für eine Erweiterung von X designiert sind. Mit dem Antilogismus verhält es sich anders. Man betrachte $I(P) = I(Q) = 1$ und $I(R) = b$. Dann ist $I(\neg R) = b$ und $I(\neg Q) = 0$. Also sind alle Werte in der Sequenz $P, Q \models R$ designiert, ebenso die Prämissen der Sequenz $P, \neg R \models \neg Q$, nicht jedoch die Konklusion.

3. Routley-Modelle

Die DeMorgan-Logik **DML** zeichnet sich durch eine erfreuliche Stabilität unter verschiedenen Perspektiven aus.

1. **DML** entsteht als natürliche Verknüpfung zweier dreiwertiger Logiken, **K3** und **LP**. Diese Logiken basieren auf den starken Kleene-Tafeln und unterscheiden sich nur in der jeweils gut motivierten Wahl designierter Werte. Die starken Kleene-Tafeln sind selbst als stärkstmögliche Definition der Wahrheitsfunktionen in drei Werten in natürlicher Weise ausgezeichnet.
2. **DML** ist die Logik tautologischer Folgerung im Sinne des oben vorgestellten Entscheidungsverfahrens. Tautologische Folgerung präzisiert die Vorstellung, nach der logische Konklusionen etwas vom Inhalt der Prämissen nach Umformungen "auspacken"; tautologische Folgerung führt das Folgern letztlich auf die Gemeinsamkeit von Variablen nach Umformung in Normalformen zurück.
3. Unser Axiomensystem für **DML** zeigt deutlich die Stelle, an der von der klassischen Logik abgewichen wird. Sie ersetzt den, aus der Sicht tautologischer Folgerung nicht akzeptablen Antilogismus durch die Regel der Kontraposition. Ansonsten bleiben alle klassischen Eigenschaften der Junktoren und der Folgerungsrelation erhalten.

Wir haben somit drei weitgehend voneinander unabhängige und natürliche Ansätze zur Beschreibung einer Folgerungsrelation, die alle die DeMorgan-Logik **DML** auszeichnen. Dem wollen wir jetzt eine vierte Beschreibung von **DML** hinzufügen.

Definition 3. Ein *Routley-Rahmen* ist ein Paar $(W, {}^*)$, welches unter den folgenden Bedingungen steht:

1. W ist eine nicht-leere Menge;
2. * ist ist eine Involution auf W (der Routley-Stern), d.h.
   $^* : W \longrightarrow W$ so, daß $a^{**} = a$ $(\forall a \in W)$.

Ein *Routley-Modell* (auf einem Routley-Rahmen) ist ein Tripel $(W, {}^*, I)$, in dem

3. I eine Interpretation der Atome in der Menge $\{0, 1\}$ ist, d.h.
   $I : \mathrm{ATM} \times W \longrightarrow \{0, 1\}$.

Wir erweitern I zu einer *Interpretation* aller Formeln (in $\{\neg, \wedge, \vee\}$) mit Hilfe einer Wahrmacherrelation $\models_I \subseteq W \times \mathrm{FML}$ wie folgt (Subskript I unterdrückt):

$$(P) \qquad a \models P \text{ gdw } I(P, a) = 1$$

(¬) $a \models \neg A$ gdw $a^* \not\models A$

(∧) $a \models A \wedge B$ gdw $a \models A$ und $a \models B$

(∨) $a \models A \vee B$ gdw $a \models A$ oder $a \models B$

Schließlich *folge* B *aus* X *im Modell* $\mathcal{M}$ ($X \models_{\mathcal{M}} B$) unter der Bedingung

$$\forall a \in W: \text{ wenn } a \models A \text{ (für alle } A \in X), \text{ dann } a \models B.$$

Die Folgerung von B aus X sei genau dann *gültig* ($X \models B$), wenn $X \models_{\mathcal{M}} B$ in beliebigen Modellen $\mathcal{M}$.

Routley-Modelle sind also zweiwertig. Sie unterscheiden sich von klassischen Modellen durch die Indexmenge W ("Punkte", "Welten") und die Stern-Operation *. Wenn wir Routley-Modelle unter die Bedingung $a = a^*$ stellen, dann kollabieren diese zu klassischen Modellen. Denn erstens wird dann jeder Punkt aufgrund der Bedingung (¬) vollständig und konsistent, d.h. (¬) mutiert zu der vertrauten Bedingung $a \models \neg A$ gdw $a \not\models A$; und zweitens verweisen dann die Bedingungen für die Interpretation einer Formel an einem Punkt niemals auf andere Punkte. Die Relativität der Interpretation zu einem Punkt dreht somit kein Rad mehr, weshalb wir sie ganz fortlassen können.

Die soeben angestellte Beobachtung erklärt, warum die relevanzlogische Negation, unter der natürlichen Annahme, daß sie eine den Wahrheitswert umkehrende Funktion ist, den Index verschieben muß. Technisch gesehen, leistet das der Routley-Stern in bestechend einfacher Weise. Aber was sollen wir uns als Indizes vorstellen und was genau soll das $*$-Bild eines solchen Index sein?

Der verwendete Buchstabe W für die Modellmenge von Indizes könnte nahelegen, die Indizes, wie in der Modallogik, als mögliche Welten zu verstehen. Mögliche Welten sind maximal konsistente Weisen, wie eine Welt aussehen könnte. Aber die Indizes der Routley-Modelle sind weder maximal (vollständig) noch konsistent. Ein Punkt kann derart sein, daß weder A noch $\neg A$ wahr ist, oder auch so, daß sowohl A also auch $\neg A$ wahr ist. Wenn wir nicht von unvollständigen oder widersprüchlichen Welten sprechen wollen, dann ist es besser, die Punkte nicht im Sinne von Welten zu verstehen. In der Relevanzlogik spricht man gern von Szenarien, Fällen (*set-ups*), Theorien oder Situationen.[9] Wir wollen uns hier weitgehend der Rede von

[9] Der Begriff der Situation ist sicher naheliegend und Mares [209], der die Frage nach der Interpretation der Indizes ausführlich behandelt, macht davon systematisch Gebrauch. Das sollte allerdings nicht das Mißverständnis einladen, daß die Semantik der Relevanzlogik eine besondere Affinität zur sogenannten Situationssemantik von Barwise [20] u.a. hat.

"Situationen" anschließen. Situationen sind einerseits so etwas wie Teile einer Welt und müssen daher nicht in jeder Hinsicht bestimmt sein. Andererseits wollen wir aber auch widersprüchliche Situationen in den Blick nehmen. Insofern widersprüchliche Situationen unmöglich sind, können sie nicht Teil einer möglichen, sondern nur Teil einer unmöglichen Welt sein.

Man könnte an dieser Stelle versuchen, die Interpretation der Punkte in der Routley-Semantik ontologisch zu dämpfen, indem nicht von Teilen der Welt (Situationen), sondern von *Beschreibungen* (Theorien) von Teilen der Welt gesprochen wird. Natürlich ist es möglich, daß Theorien von Teilen einer möglichen Welt widersprüchlich sind. Aber wenn wir die Sätze einer Sprache relativ zu einer Theorie T interpretieren, dann erhalten wir Bedingungen dafür, wann ein Satz T zufolge *behauptbar* ist. Die These, daß Behauptbarkeits- und nicht Wahrheitsbedingungen die Bedeutung der Junktoren bestimmen, ist aber sicher nicht Teil des relevanzlogischen Unternehmens. Es wäre deshalb nicht "im Sinne der Erfinders", die Interpretation der Routley-Semantik unter die Bedingung dieser These zu stellen. Wir lassen hier die weitere Erörterung der philosophischen Interpretation der Indexmenge W offen und verweisen auf die Diskussion in [24, 26, 54, 209, 227, 228].

Zum Routley-Stern gibt es verschiedene heuristische Erläuterungen. Vielleicht die beste ist die folgende.[10] Dazu brauchen wir drei Begriffe:

· Situationen können miteinander *verträglich* sein, in dem Sinne, daß sie einander logisch nicht ausschließen.

· Eine Situation kann *Teil* einer anderen sein: was wahr in einer Situation ist, ist in jeder ihrer Teilsituationen wahr. (Wir schreiben $a \leq b$, wenn a Teil von b ist.)

· Eine Situation kann sich zu einer anderen *spiegelbildlich* verhalten. Was unter dem Spiegelbild a^* einer Situation a zu verstehen ist, drücken wir am besten so aus:

$$A \text{ in } a \text{ gdw } \neg A \text{ nicht in } a^*,$$
$$A \text{ in } a^* \text{ gdw } \neg A \text{ nicht in } a.$$

Daraus folgt unmittelbar, daß $a = a^{**}$. Ebenso: Eine Situation ist genau dann unvollständig in Bezug auf eine Formel, wenn dessen Spiegelbild widersprüchlich bezüglich dieser Formel ist.

[10] Sie knüpft an die Interpretation der Negation in der intuitionistischen Logik an; siehe Kap. 2.9 und vgl. Goldblatt [116]. In Routley *et al.* [254] wird sie kurz angerissen; Dunn [67] und Restall [246, 248] führen den unten skizzierten Zusammenhang zwischen dem Stern und der Kompatibilitätsrelation systematisch aus. Eine kurze Darstellung findet sich auch in Mares [209, Kap. 5.2].

Nun ist es sicher so, daß wenn $\neg A$ in einer Situation wahr ist, dann gibt es keine mit dieser verträglichen Situation, in der A wahr ist, und umgekehrt. Wenn wir Cab für "a ist mit b verträglich" schreiben, dann haben wir gerade die folgende Wahrheitsbedingung für $\neg A$ beschrieben:

$$a \models \neg A \text{ gdw } \forall b : Cab \Rightarrow b \not\models A. \tag{72}$$

Ferner gilt: Wenn a mit b verträglich ist und A in b wahr ist, dann muß $\neg A$ in a falsch sein. Aber $\neg A$ ist genau dann in a falsch, wenn A im Spiegelbild a^* von a wahr ist. Wenn also Cab, dann muß b ein Teil von a^* sein.[11] Da das Ganze auch umgekehrt gilt, haben wir

$$Cab \text{ gdw } b \leq a^*.$$

Das ergibt zusammen mit (72)

$$a \models \neg A \text{ gdw } \forall b : b \leq a^* \Rightarrow b \not\models A. \tag{73}$$

Die rechte Seite von (73) ist aber äquivalent zur rechten Seite von ($\neg$). Das folgt aus der oben angegebenen Erklärung der Teilrelation $\leq$, für die wir hier nur annehmen müssen:

Reflexivität. $\qquad a \leq a;$

Weitergabe. $\qquad a \leq b \;\&\; a \models A \Rightarrow b \models B.$

Aufgrund der Reflexivität, folgt aus (73)-Rechts, daß $a^* \not\models A$. In umgekehrter Richtung nehmen wir an, daß $a^* \not\models A$ sowie $b \leq a^*$. Dann folgt aus der Bedingung der Weitergabe $b \not\models A$, wie gewünscht. So haben wir gezeigt, daß die zunächst intuitiv untermotiviert anmutende Negationsbedingung ($\neg$) eine besonders elegante Formulierung der gut motivierten Bedingung (72) ist.

In die Erklärung der Spiegelbildlichkeit haben wir die Involutionsbedingung $a = a^{**}$ gleich eingebaut. An dieser Bedingung hängt die Äquivalenz der doppelten Verneinung mit der Bejahung. Im Bild ist das natürlich richtig: Ein Bild zweimal gespiegelt ergibt das Originalbild. Aber die Frage, wie wir mit doppelten Verneinungen logisch umgehen sollen, können wir sicher nicht aus den geometrisch-optischen Gesetzen für Spiegel beantworten. Wünschenswert wäre daher eine Rückführung von ($\neg$) auf (72), in der wir nicht von der Spiegelbildlichkeit Gebrauch machen, d.h. in der auch die Involutionsbedingung für den Stern aus Eigenschaften der Kompatibilitätsrelation folgt. Restall [248] zeigt, wie dieser Wunsch erfüllt werden kann.[12]

[11] Natürlich gilt auch, daß wenn Cab, dann $a \leq b^*$, denn Verträglichkeit (C) ist plausiblerweise eine symmetrische Relation.

[12] Dort werden die oben angestellten Überlegungen über die Teil- und Verträglichkeitsbeziehung auch formal präzise durchgeführt. Dabei sind einige Details zu beachten, die den Rahmen unserer Darstellung sprengen würden.

Routley-Modelle und der DeMorgan-Diamant. Es sei $\mathbb{R}$ die Klasse der Routley-Modelle. Wir wollen nun zeigen, daß eine Sequenz $X \models A$ genau dann in $\mathbb{R}$ gültig ist, wenn sie im DeMorgan-Diamanten gültig ist.

In Routley-Modellen führt eine zweifache Anwendung des Sterns auf eine Situation zu dieser Situation zurück (denn $a = a^{**}$). Das bedeutet, daß fortgesetzte Anwendungen der Stern-Funktion immer nur zwischen zwei Punkten hin- und herspringen. Daher genügt es, im Sinne des folgenden Lemmas, Routley-Modelle mit zwei Elementen zu betrachten.

Lemma 4. *Eine Sequenz $X \models A$ ist genau dann in der Klasse $\mathbb{R}$ aller Routley-Modelle gültig, wenn sie in der Klasse $\mathbb{R}^2$ der zwei-elementigen Routley-Modelle $(\{a, b\},^*, I)$ $(a \neq b)$ gültig ist.*

Beweis. (Im wesentlichen verfahren wir so wie im Beweis von Satz 32 über erzeugte Teilmodelle im Kapitel über Modallogik.) Da $\mathbb{R}^2 \subseteq \mathbb{R}$ gilt das Lemma von links nach rechts.

Angenommen nun, $X \models A$ gilt nicht in $\mathbb{R}$; d.h. es gibt einen Punkt a in einem Modell $\mathcal{M} \in \mathbb{R}$ so, daß $a \models X$ und $a \not\models A$. Wir zeigen, daß sich am Punkte a ein Teilmodell $\mathcal{M}^a = (W^a,^*, I^a)$ von $\mathcal{M}$ erzeugen läßt, in welchem $X \models A$ ebenfalls widerlegt ist.

Es sei Stern(a) der transitive Abschluß der Menge $\{a\}$ unter der Sternoperation. Da die Operation involutiv ist, besteht Stern(a) aus nur zwei Elementen. Das Teilmodell $\mathcal{M}^a$ – mit * wie in $\mathcal{M}$ – sei nun so definiert:

$$W^a = \text{Stern}(a) = \{a, a^*\};$$

$$\forall x \in W^a: \ I^a(P, x) = 1 \text{ gdw } I(P, x) = 1 \text{ in } \mathcal{M} \ (P \in \text{ATM}).$$

Die Interpretation I^a in $\mathcal{M}^a$ erweitern wir zu einer Interpretation aller Formeln unter den Bedingungen $(\neg)$, $(\wedge)$ und $(\vee)$, wie in $\mathbb{R}$. $\mathcal{M}^a$ ist offensichtlich ein Routley-Modell mit nur zwei Elementen. Durch Induktion über den Formelaufbau läßt sich nun zeigen, daß eine Formel in $\mathcal{M}$ genau dann gilt, wenn sie in $\mathcal{M}^a$ gilt. A, so daß $a \models A$ in $\mathcal{M}$ genau dann, wenn $a \models A$ in $\mathcal{M}^a$. Der Fall, daß eine Formel A ein Atom ist, ist durch die Definition von I^a gegeben. Im Fall $A = \neg B$ argumentieren wir so:

$$\begin{aligned} a \not\models \neg B \text{ in } \mathcal{M} \ &\text{gdw } a^* \models B \text{ in } \mathcal{M}, \ \text{nach } (\neg) \\ &\text{gdw } a^* \models B \ \text{ in } \mathcal{M}^a, \ \text{nach Induktionsannahme} \\ &\text{gdw } a \not\models \neg B \text{ in } \mathcal{M}^a, \ \text{nach } (\neg). \end{aligned}$$

Die RL-Richtung des Lemmas folgt nun unmittelbar. ∎

Eine Semantik hat die Aufgabe, Sätzen eine Bedeutung zuzuordnen. Wie schon zuvor, wollen wir sagen, daß die Bedeutung eines Satzes – sein semantischer Wert – die *Proposition* ist, die er ausdrückt. Je nach semantischer Theorie fällt der Propositionsbegriff mehr oder weniger abstrakt aus. So betrachten wir in der klassischen Aussagenlogik nur zwei semantische Werte, die Sätze annehmen können: Wahrheit und Falschheit, 0 und 1. Da in wahrheitsfunktionalen Logiken generell Sätze auf Wahrheitswerte abgebildet werden, sind darin höchstens soviele logisch distinkte Propositionen ausdrückbar wie es distinkte Wahrheitswerte gibt: in **KL** zwei, in **L3** und **LP** drei, und in **DML** vier. In Modellen mit einer Punktmenge W, wie wir sie aus der Modallogik kennen, sind Propositionen etwas komplexer. Wir können sie auffassen als Funktionen von W in die angenommene Menge der Wahrheitswerte, in der Regel die Werte 0 und 1:

$$\text{Proposition} : \mathrm{W} \longrightarrow \{0, 1\}.$$

Danach sind Propositionen Paare in der Menge $W \times \{0, 1\}$. Wir können solche Paarmengen auch einfacher darstellen, nämlich als die Menge der Punkte, deren zweite Koordinate der Wert 1 ist. D.h. wenn ϕ ein Proposition im Funktionssinne ist, dann können wir ebensogut von der Menge

$$|\phi| = \{x \in W : \phi(x) = 1\}$$

als einer Proposition im abgeleiteten Sinne sprechen.

In einem Modell mit einer Punktmenge W gibt es im Prinzip so viele Propositionen wie es Mengen von Punkten, d.h. Elemente in $\wp(W)$ gibt.[13] Auch die zweiwertigen Routley-Modelle basieren auf Punktmengen. Nach dem Lemma wissen wir, daß wir uns für die Frage der Gültigkeit in der Klasse der Routley-Modelle auf die Betrachtung zwei-elementiger Modelle beschränken können. In diesen gibt es nur genau vier ausdrückbare Propositionen, die wir durch die Elemente der Menge $\wp(\{a, a^*\})$ darstellen können:

$$\{a, a^*\}, \{a\}, \{a^*\}, \emptyset.$$

In $\mathbb{R}^2$-Modellen gibt es also genauso viele semantische Werte wie im DeMorgan-Diamanten.

Mehr noch, diese vier Propositionen sind genauso "angeordnet" (siehe gleich) wie im DeMorgan-Diamanten. Die Grundlage dieser Beobachtung

[13] Im Prinzip: Das soll nicht ausschließen, daß wir für bestimmte Zwecke, die Menge der Propositionen in einer Punktmenge unter Bedingungen stellen.

ist ein sehr allgemeines Verfahren, welches Kripke [165] darstellt und – etwas verkürzt – so zusammengefaßt werden kann:

- Zu jeder Klasse endlicher Kripke-Rahmen gibt es eine endlich-wertige Matrix so, daß Gültigkeit in der Rahmenklasse und Gültigkeit in der Matrix übereinstimmen. [165, cf. p. 93]

Routley-Rahmen sind Kripke-Rahmen genügend ähnlich. Wir zeichnen nun das Verfahren für den uns interessierenden Fall nach. Man nehme die Menge $M = \wp(\{a, a^*\})$. Es seien α und β Propositionen aus der Grundmenge $W = \{a, a^*\}$ eines Routley^2-Rahmens gebildet. Dann definieren wir drei Operationen auf Propositionen, für alle $x \in W$:

$$\text{(Op)} \qquad \begin{aligned} \neg\alpha(x) = 1 &\text{ gdw } \alpha(x^*) \neq 1; \\ \alpha \wedge \beta(x) = 1 &\text{ gdw } \alpha(x) = \beta(x) = 1; \\ \alpha \vee \beta(x) = 1 &\text{ gdw } \alpha(x) = 1 \text{ oder } \beta(x) = 1. \end{aligned}$$

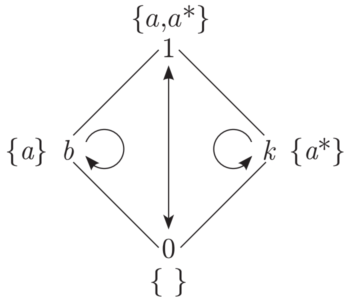
Ohne Gefahr einer Verwechslung sind die Operationen hier durch die Junktorzeichen bezeichnet. Mit der abgeleiteten Auffassung von Propositionen können wir das auch so ausdrücken ($X^* = \{x^* : x \in X\}$):

$$|\text{Op}| \qquad \begin{aligned} |\neg\alpha| &= W \setminus |\alpha|^*; \\ |\alpha \wedge \beta| &= |\alpha| \cap |\beta|; \\ |\alpha \vee \beta| &= |\alpha| \cup |\beta|. \end{aligned}$$

Was sagen diese Äquivalenzen bzw. Gleichungen aus über das Resultat der Operationen in Abhängigkeit von ihren Argumenten? Nehmen wir den Fall der Negationsoperation $\neg$. Die Proposition $|\alpha|$ kann (a) $\{a, a^*\}$, (b) $\{a\}$, (c) $\{a^*\}$, oder (d) $\{\,\}$ sein. Wenn (a), dann ist $|\alpha|^* = \{a, a^*\}$ und also ist $W \setminus |\alpha|^* = \emptyset$ – und umgekehrt. Wenn (b), dann ist $|\alpha|^* = \{a^*\}$ und also ist $W \setminus |\alpha|^* = \{a\}$ – und umgekehrt. Der Fall (c) gilt genau dann, wenn $W \setminus |\alpha|^* = \{a^*\}$, und der Fall (d) genau dann, wenn $W \setminus |\alpha|^* = \{a, a^*\}$. Wir können das in einer Tafel zusammenfassen:

α	$\neg\alpha$
$\{a, a^*\}$	$\{\,\}$
$\{a\}$	$\{a\}$
$\{a^*\}$	$\{a^*\}$
$\{\,\}$	$\{a, a^*\}$

Ein Vergleich mit den **DML**-Tafeln auf p. 341 zeigt, daß das genau die Tafel für die Negation ist, in der die Werte 1, b, k, und 0 nun durch die Propositionen $\{a, a^*\}$, $\{a\}$, $\{a^*\}$, und $\{\,\}$ vertreten werden. Auf die gleiche Weise erhalten wir die **DML**-Tafeln für $\wedge$ und $\vee$, in denen die Werte ebenso neu bezeichnet sind. Wir haben aus einem zwei-elementigen Routley-Rahmen den DeMorgan-Diamanten erzeugt:



Das ist die Schlüsselbeobachtung, aus der die Äquivalenz von zwei-elementigen Routley-Modellen und dem DeMorgan-Diamanten hergeleitet werden kann. Die Details findet der Leser in [254, §2].

Als Überleitung zum nächsten Abschnitt diene die folgende

Übung. Man betrachte den Junktor $\supset$ mit $A \supset B := \neg A \vee B$ und untersuche seine Eigenschaften in **DML** im Hinblick auf die Frage, ob es sich dabei um einen guten Implikationsoperator handeln kann – aufgrund allgemeiner Überlegungen über die Eigenschaften der Implikation, aber auch aus parakonsistenter bzw. relevanzlogischer Sicht.

4. Theorien der Implikation

Die Logik **DML** definiert zwar eine relevanzlogische Folgerungsbeziehung – jedenfalls im Sinne der Überschneidungsbedingung –, jedoch keinen relevanzlogischen Implikationsoperator. Dem steht eine prinzipielle Schwierigkeit entgegen. Angenommen, wir fügen den vierwertigen Tafeln für $\neg$, $\wedge$ und $\vee$ eine Tafel für $\to$ hinzu. Dann verliert $\models$ die Eigenschaft eine relevanzlogische Folgerungsbeziehung zu sein, sobald diese Tafel eine Implikationsformel als gültig auszeichnet. Sei zum Beispiel die Tafel für $\to$ so, daß $A \to A$ gültig ist, d.h. für jeden Wert von A einen designierten Wert annimmt. Dann folgt aus der Definition (p. 342) einer gültigen Folgerung in **DML**, daß $X \models A \to A$ für eine beliebige Wahl von X; denn $A \to A$ hat ja einen designierten Wert, gleichgültig welche Werte die Formeln in X haben. So hätten wir zum Beispiel die Instanz $Q \models P \to P$, welche die Bedingung der Variablenüberschneidung offensichtlich nicht erfüllt. Die Überlegung zeigt übrigens auch, daß die Relation $\models$ in **DML** nur deshalb die Überschneidungsbedingung erfüllt, weil es in **DML** weder gültige Formeln gibt, noch solche, die nie einen designierten Wert erhalten. So ist insbesondere $A \vee \neg A$ nicht gültig, noch nimmt $A \wedge \neg A$ unter jedem Wert von A einen nicht-designierten Wert an. Die intuitive Interpretation der Werte begründet das gut: Eine Formel kann halt weder wahr noch falsch sein (k) oder sie kann beides sein (b). Im ersten Fall ist $A \vee \neg A = k$ (also undesigniert), im zweiten Fall ist $A \wedge \neg A = b$ (also designiert). Aber wie wollte man auf ähnlich gute Weise begründen, daß $A \to A$ nicht gültig sein kann? Wir sehen, der mit **DML** eingeschlagene Weg läßt sich nicht einfach fortsetzen, wenn uns an einer iterierbaren, relevanten Implikation gelegen ist.[14]

Dennoch können wir **DML** als *Grundlegung* einer Theorie der relevanten Implikation auffassen. Wenn wir nämlich gültige Sequenzen der Form $A \models B$ betrachten und diese in Formeln $A \to_0 B$ übersetzen, dann erhalten wir eine Menge $\mathrm{DML}^{\to_0}$ erststufiger Implikationsformeln, welche jede Theorie iterierbarer, relevanter Implikation konservativ erweitern sollte. Eine *bona fide* Relevanzlogik sollte also die Bedingung erfüllen, daß sie $\mathrm{DML}^{\to_0}$ einschließt und $\mathrm{DML}^{\to_0}$ keine weiteren erstufigen Formeln hinzufügt. Etwas anders ausgedrückt: Jede Relevanzlogik im vollen Vokabular sollte so sein, daß die Tafeln für **DML** ein Entscheidungsverfahren für erststufige Implikationen darstellen.

Dies ist einer von zwei sich ergänzenden Ansätzen, die charakteristisch für die Entwicklung der Relevanzlogik sind. Der andere Ansatz besteht darin, zunächst eine reine Theorie der Implikation zu entwerfen. Diese Theorie soll dann um weitere Junktoren erweitert werden, wobei $\mathrm{DML}^{\to_0}$ diese Erweiterungen kontrolliert. Diesem zweiten Ansatz werden wir uns im folgenden zuwenden.

* * *

Es ist klar, daß ein rein implikatives Schema wie

$$A \to (B \to B) \tag{74}$$

nicht zum Theorembestand einer Relevanzlogik gehören kann. Das Schema hat zuviele Instanzen, die den Test der Variablenüberschneidung nicht bestehen. Das gilt zwar nicht für

$$A \to (B \to A). \tag{75}$$

[14] In einer Fußnote auf p. 336 haben wir Makinson zitiert: “In the author’s opinion, in the realm of nonclassical logic where nothing has usefulness to recommend it, the only justification for existence is beauty, ...” [193, p. 41]. Makinson setzt den Satz so fort: “... and systems transforming DeMorgan implication into an operator lack it sadly.”

Aber (74) und (75) sind logisch äquivalent (nach Umbuchstabierung), wenn einfache Prämissenumstellung

C. $$(A \to (B \to C)) \to (B \to (A \to C))$$

erlaubt ist – und das ist sie in den meisten Relevanzlogiken. Wie können wir also Schemata wie (74) und (75) als ungültig ausschließen?

Vier (erfolglose) Versuche. Klassische Modelle der Implikation sind (trivial) einpunktige Modelle $(\{a\}, I)$, in denen die Funktion I Wahrheitswerte über Atome am Punkt a verteilt. (Da die Indexmenge $\{a\}$ keine Verteilungen unterscheidet, ist sie redundant.) Die Bedingung für die (materiale) Implikation ist

(76) $$a \models A \to B \text{ gdw } a \models A \Rightarrow a \models B.$$

Hier, wie auch im folgenden, kürzt $\Rightarrow$ die metasprachliche materiale Implikation ab. D.h. $a \models A \Rightarrow a \models B$ bedeutet soviel wie:

> nicht: $a \models A$ und $a \not\models B$; d.h.
> entweder $a \not\models A$ oder $a \models B$.

Es ist unmittelbar klar, daß (75) unter dieser Bedingung gültig sein muß.

Im zweiten Versuch fassen wir den Pfeil im Sinne einer universalen strikten Implikation auf. D.h. wir betrachten Modelle (W, I), in denen W $(\neq \emptyset)$ nun mehr als nur einen Punkt enthalten kann und interpretieren den Pfeil so:

(77) $$a \models A \to B \text{ gdw } \forall b : b \models A \Rightarrow b \models B.$$

Wenn B eine gültige Formel ist – wie $C \to C$ –, dann muß $A \to B$ gültig sein. So wird also das unerwünschte Schema (74) gültig.

Daran ändert sich auch im dritten Versuch nichts, in dem wir eine zweistellige Relation R $(\subseteq W \times W)$ in die Modelle einführen und den Pfeil unter die folgende Bedingung stellen:

(78) $$a \models A \to B \text{ gdw } \forall b : Rab \;\&\; b \models A \Rightarrow b \models B.$$

Die Bedingung definiert die strikte Implikation in relationalen Modellen der Modallogik (Kripke-Modelle). Die strikte Implikation ist jedoch nichts weiter als eine modalisierte materiale Implikaton (die wir vorübergehend mit $\supset$ notieren wollen). Das wird sofort deutlich, wenn wir (78) äquivalent umschreiben zu:

(79) $$a \models A \to B \text{ gdw } \forall b : Rab \Rightarrow (b \not\models A \text{ oder } b \models B).$$

Der Ausdruck in Klammern auf der rechten Seite ist natürlich die Wahrheitsbedingung für die materiale Implikation $A \supset B$ am Punkt b. Nun werden aber klassische Tautologien an allen Punkten wahr. Das gilt insbesondere auch für tautologische Formeln der Form $A \supset B$. Wenn $A \supset B$ tautologisch ist, dann ist $A \to B$ nach (79) an allen Punkten, d.h. logisch wahr. Es ist klar, daß aus relevanzlogischer Sicht die strikte Implikation kaum besser als die materiale ist.

Das Problem in den bisherigen Versuchen ergab sich aus dem Umstand, daß es keine Punkte gibt, an denen logisch wahre Formeln falsch werden können. Im vierten Versuch führen wir daher genau solche Punkte ein. Die Modelle $(W, 0, I)$ bestehen nun aus einer Menge W von Punkten, einschließlich solchen, an denen Formeln wie $P \to P$ falsch sein können; einem normalen Punkt 0 (an dem $P \to P$ immer wahr ist); und einer Interpretation I, die nun Wahrheitswerte nicht nur über Atome, sondern über *alle* Formeln verteilt. So kann eine Bewertung I der Formel $P \to P$ an einem Punkt in W einfach den Wert 0 zuordnen, gleichgültig, welchen Wert P an diesem Punkt hat.[15] Der normale Punkt 0 ist jedoch nicht von dieser Art. Dieser steht vielmehr unter der Bedingung

(80) $$0 \models A \to B \text{ gdw } \forall a \in W : a \models A \ \Rightarrow \ a \models B.$$

Eine Formel ist wahr in einem solchen Modell, wenn sie am 0-Punkt des Modells wahr ist. Gültig sind solche Formeln, die in allen Modellen wahr sind.

Wenn wir jetzt eine Instanz von (74) betrachten, z.B. $P \to (Q \to Q)$, dann erlaubt die Bedingung ein Gegenmodell wie das folgende:

- Der Punkt a sei nicht normal. Dann können wir eine Bewertung I wählen so, daß $I(P, a) = 1$, jedoch $I(Q \to Q, a) = 0$. In diesem Fall kann $P \to (Q \to Q)$ am 0-Punkt nicht wahr sein.
- Dennoch ist $Q \to Q$ eine gültige Formel. Denn am 0-Punkt muß auch $Q \to Q$ unter der rekursiven Interpretationsbedingung für Implikationen stehen.

Diese Theorie der Implikation ist das implikative Fragment $\mathbf{N4}^{\to}$ der Logik **N4** (siehe Priest [227, 228, Kap. 9].) Keine der relevanzlogisch abzulehnenden Implikationsformeln ist ein Theorem von $\mathbf{N4}^{\to}$. Der Preis ist jedoch hoch. Denn da die Interpretation von Pfeilformeln, außer in 0, unter keinen

[15] Obacht: Wir verwenden hier, um nicht unnötig von der üblichen Notation in der Literatur abzuweichen, “0” sowohl um den normalen Punkt eines Modells als auch um einen der zwei Wahrheitswerte zu bezeichnen.

rekursiven Bedingungen steht, kann $\mathbf{N4}^{\rightarrow}$ mit verschachtelten Implikationen, wie dem Schema C der Prämissenvertauschung, nichts anfangen. Letztlich sind nur Instanzen des Identitätsschemas $A \rightarrow A$ Theoreme. Man beweise die folgende Beobachtung als Übung.

BEOBACHTUNG 5. *$A \rightarrow B$ ist genau dann ein Theorem von $\mathbf{N4}^{\rightarrow}$, wenn $A = B$.*

Für die volle Logik **N4** gilt die Beobachtung nicht, da Negationen, Konjunktionen und Disjunktionen an allen Punkten unter rekursiven Bedingungen stehen.[16] Aber auch in **N4** kommen wir so nur um *eine* gehaltvolle Implikationsstufe über die DeMorgan-Logik hinaus. Auch dieser vierte Versuch bringt uns daher nicht weiter in unserer Suche nach einer interessanten Theorie beliebig verschachtelter Implikationen.

Dreistellige Zugangsrelation (Routley-Meyer-Semantik). Implikation durch eine zweistellige Relation zu modellieren, so haben wir im dritten Versuch gesehen, führt nicht zu den gewünschten Resultaten. Zweistellige Relationen ergeben einfach nicht die für eine gute Theorie der Implikation erforderliche Modellierungsbreite: Die Modelle sind so, daß unerwünschte Formeln wie (74) nicht aus dem Kreis logisch wahrer Implikationen ausgeschlossen werden können.

Nach der Zahl 2 kommt die Zahl 3. Wenn zweistellige Relationen nicht zum Ziel führen, dann ist es völlig naheliegend, es mit dreistelligen Relationen und der folgenden Bedingung zu versuchen:

$$(\rightarrow) \qquad a \models A \rightarrow B \text{ gdw } \forall bc : Rabc \;\&\; b \models A \;\Rightarrow\; c \models B.$$

Wenn jetzt die angestrebte Modellierungsbreite erzielt wird, dann ist diese semantische Theorie eindeutig besser als jeder der bisher betrachteten Versuche. Und da es für den gesetzten Zweck unsinnig wäre, vierstellige Relationen mit noch weiterer Modellierungsbreite in den Blick zu nehmen, ist die Theorie mit einer dreistelligen Relation genau richtig. Die dreistellige Relation ist ein theoretischer Term in einer semantischen Theorie der Implikation. Dieser Term ist in dem Maße richtig gewählt, in dem die semantische Theorie erfolgreich die Grundeigenschaften einer Implikation aus relevanzlogischer Perspektive darstellt.

Manche Kommentatoren der Routley-Meyer-Semantik (z.B. Copeland in [54]) beklagen, daß es keine über das bisher gesagte hinausgehende "In-

[16] Die Konjunktion und Disjunktion stehen in **N4** unter den Boole'schen Bedingungen. Für die Negation werden die Rahmen um den Routley-Stern erweitert und die entsprechende Interpretationsbedingung erweitert. Details sind in [227, 228] nachzulesen.

terpretation” der dreistelligen Relation gibt, daß diese Semantik “bloß formal” sei.[17] (Ähnliches wird gelegentlich auch über die Kripke-Semantik der Modallogik gesagt.) Vielleicht deutet sich in solcher Kritik eine tiefe Kontroverse über die Aufgabe einer semantischen Theorie an, vielleicht handelt es sich auch nur um ein Mißverständnis über das Wesen theoretischer Terme im allgemeinen und in semantischen Theorien im besonderen. Wie dem auch sei, wir werden später (pp. 375f.) vorführen, wie die dreistellige Relation sich in ein aus der Konditonallogik bekanntes und relativ unkontroverses Schema einfügt. In diesem Sinne werden wir eine indirekte Interpretation der dreistelligen Relation anbieten.

Wir beginnen damit, die Rahmen und Modelle der Routley-Meyer Semantik in einer Definition vorzustellen.

Definition 6. (A-Rahmen und Modelle.)
Ein *A-Rahmen* ist ein Tripel $(W, 0, R)$ so, daß

1. W eine nichtleere Menge ist;
2. $0 \in W$ (die “aktuale Situation” des Modells);
3. $R \subseteq W^3$ eine dreistellige Relation (“Zugang”) zwischen Situationen ist.

Drei Abkürzungen sind jetzt nützlich:

$$\begin{aligned} a \le b :&\ \text{gdw } R0ab \\ R^2(ab)cd :&\ \text{gdw } \exists x : Rabx \ \&\ Rxcd \\ R^2a(bc)d :&\ \text{gdw } \exists x : Raxd \ \&\ Rbcx \end{aligned}$$

Alle Rahmen stehen unter den folgenden Bedingungen:

(R1)	Reflexivität	$a \le a$;
(R2)	Monotonie	$a \le b \ \&\ Rbcd \ \Rightarrow\ Racd$;
(R3)	Assoziativität	$R^2(ab)cd \ \Rightarrow\ R^2a(bc)d$;
(R3*)	Assoziativität*	$R^2(ab)cd \ \Rightarrow\ R^2b(ac)d$.

Ein *Modell* $(W, 0, R, v)$ auf einem A-Rahmen fügt diesem eine Interpretation $I : \mathrm{ATM} \longrightarrow \{0, 1\}$ hinzu unter der Bedingung:

(Atm $\le$) Atomare Weitergabe $\quad I(P, a) = 1 \ \&\ a \le b \ \Rightarrow\ I(P, b) = 1.$

Ein Modell wird zu einem *Modell für* $\rightarrow$ erweitert durch die Definition einer Wahrmacherrelation $\models\ \subseteq W \times \mathrm{FML}$ für alle Formeln, welche unter den Bedingungen $a \models P$ gdw $I(P, a) = 1$ (für alle $P \in \mathrm{ATM}$) sowie

[17] Siehe dazu [26], sowie [24].

$$(\rightarrow) \qquad a \models A \rightarrow B \text{ gdw } \forall bc : Rabc \;\&\; b \models A \;\Rightarrow\; c \models B$$

steht. (*Ende der Definition.*)

Wenn wir von Modellen sprechen, dann meinen wir in diesem Abschnitt immer Modelle für $\rightarrow$. Die Definitionen von Wahrheit und Folgerung sind wieder gestaffelt nach den Modell- und Rahmenmengen, über die wir quantifizieren: (a) ein bestimmtes Modell auf einem bestimmten Rahmen; alle Modelle auf einem bestimmten Rahmen; (b) alle Modelle auf allen Rahmen, die bestimmte Eigenschaften erfüllen.

Definition 7. (Logische Wahrheit und Folgerung.) Es sei A eine Formel und X eine Menge von Formeln.

1. (a) A ist genau dann *wahr in einem Modell* $\mathcal{M}$, wenn $0 \models A$ in $\mathcal{M}$ (Notation: $\mathcal{M} \models A$). A ist genau dann gültig in einem Rahmen $\mathcal{R}$, wenn A in allen Modellen auf $\mathcal{R}$ wahr ist ($\mathcal{R} \models A$).
   (b) A ist genau dann *gültig in einer Rahmenklasse* $\mathbb{R}$, wenn A in allen Rahmen in $\mathbb{R}$ gültig ist ($\mathbb{R} \models A$).
2. (a) A *folgt* genau dann aus X *in einem Modell* $\mathcal{M}$, wenn unter der Annahme, daß alle Formeln in X in $\mathcal{M}$ wahr sind, auch A in $\mathcal{M}$ wahr ist ($X \models_{\mathcal{M}} A$).
   (b) Die *Folgerung* aus X auf A ist genau dann *gültig in einer Rahmenklasse* $\mathbb{R}$, wenn A aus X in allen Modellen auf allen Rahmen in $\mathbb{R}$ folgt ($X \models_{\mathbb{R}} A$).

Wenn das tief- oder vorangestellte $\mathbb{R}$ fehlt, dann ist Gültigkeit bzw. Folgerung in der Klasse aller A-Rahmen (i.S.d. Definition) gemeint.

Anmerkung. Im zweiten Teil der Definition sehen wir eine klassisch bestimmte Folgerungsrelation am Werk, d.h. "wenn $\mathcal{M} \models X$, dann $\mathcal{M} \models A$" ist im Sinne von "$\mathcal{M} \not\models X$ oder $\mathcal{M} \models A$" zu verstehen. Somit ist klar, daß $X \models_{\mathcal{M}} A$ *kein* relevantes Folgerungsverhältnis ausdrückt. Es gilt beispielsweise $A \models_{\mathcal{M}} B \rightarrow B$ in jedem Modell $\mathcal{M}$ (siehe unten). Das sollte uns nicht weiter beunruhigen, denn der Zielpunkt der jetzigen Modellierung ist nicht eine relevante Folgerungsrelation, sondern eine relevante Implikation. Das Interesse richtet sich daher nicht auf Teil 2 der Definition, d.h. auf die Menge der gültigen Folgerungen, sondern auf Teil 1, d.h. die Menge der gültigen Pfeilformeln.

Daß die in Teil 2 definierte Relation nicht zur relevanten Implikation "paßt" zeigt sich schon darin, daß Modelle nicht allgemein die Äquivalenz

$$(81) \qquad A \models_{\mathcal{M}} B \text{ gdw } \models_{\mathcal{M}} A \rightarrow B$$

erfüllen. Würde (81) gelten, dann könnten wir $B \to B$ für B einsetzen und erhielten – von links nach rechts – (74) $A \to (B \to B)$ als gültige Formel. Das Schema (74) ist jedoch nicht gültig, wie wir gleich sehen werden.

Der Nullpunkt eines Modells gibt an, "wie es ist" (dem Modell zufolge). *Wahrheit im Modell* ist daher Wahrheit am Nullpunkt des Modells. Das gilt insbesondere auch für Implikationen. Nun sollte eine Implikation $A \to B$ in einem Modell genau dann wahr sein, wenn die Wahrheit von A an einem Punkt immer die Wahrheit von B erzwingt, d.h. wenn es in diesem Modell keine A-Punkte gibt, die nicht auch B-Punkte sind. Es ist die Aufgabe der Weitergabebedingung (Atm $\leq$) – zusammen mit (R1), der Reflexivität von $\leq$ – dies zunächst einmal für Atome zu garantieren. Wir schreiben wieder $[\![A]\!]$ für die Menge der Punkte, an denen eine Formel A wahr ist.

BEOBACHTUNG 8. *In einem Modell gilt* $0 \models P \to Q$ $(P, Q \in ATM)$ *genau dann, wenn* $[\![P]\!] \subseteq [\![Q]\!]$.

BEWEIS. Angenommen, (1) $0 \models P \to Q$ und (2) $a \models P$. Aus (1) nach $(\to)$ folgt, (3) $\forall xy : x \leq y \,\&\, x \models P \;\Rightarrow\; y \models Q$. Dann folgt aus (R1), (2) und (3) das gewünschte $a \models Q$.
Umgekehrt nehmen wir an, daß (1) $[\![P]\!] \subseteq [\![Q]\!]$, sowie (2) $a \leq b$ und (3) $a \models P$. Aus (1) und (3) folgt (4) $a \models Q$. Aus (2) und (4) folgt nach (R3), daß $b \models Q$. ∎

Was für Implikationen zwischen Atomen gilt, sollte ganz allgemein für beliebige Formeln gelten. Inspektion des gerade geführten Beweises zeigt, daß die Beobachtung sich in diesem Sinne verallgemeinern läßt, vorausgesetzt, die Weitergabe-Bedingung gilt nicht nur für Atome. Die Monotonie-Bedingung (R2) stellt sicher, daß diese Voraussetzung zutrifft.

LEMMA 9. (Weitergabe) *Für alle Formeln A und beliebige Punkte a in beliebigen Modellen gilt: Wenn $a \models A$ und $a \leq b$, dann $a \models B$.*

BEWEIS. Induktion über den Aufbau von A. Die *Basis* (für atomare Formeln) ist durch (Atm $\leq$) in Definition 6 gegeben. Da wir hier nur Modelle für $\to$ betrachten, ist der *Induktionsschritt* mit dem Fall $A = B \to C$ schon beendet. Wir nehmen also an, daß $a \models B \to C$, d.h. (1) $\forall xy : Raxy \,\&\, x \models B \;\Rightarrow\; y \models C$, und (2) $a \leq b$. Zu zeigen ist $b \models B \to C$. Wir nehmen also ferner an, daß (3) $Rbcd$ sowie (4) $c \models B$ (und zeigen $d \models C$.) Aus (2) und (3) folgt nach (R2) in Definition 6, daß (5) $Racd$, woraus zusammen mit (4) nach (1) das gewünschte Resultat folgt.
(Bei Erweiterungen der Sprachen werden an dieser Stelle weitere Fälle im Induktionsschritt zu betrachten sein.) ∎

KOROLLAR 10. *Für alle Formeln A und B: $0 \models A \to B$ genau dann, wenn $[\![A]\!] \subseteq [\![B]\!]$.*

Das Korollar bestätigt nicht nur, daß die Wahrheit einer Implikation in einem Modell, das Verhältnis von Antezedens- und Konsequenspunkten im Modell so reflektiert, wie wir das erwarten. Es erleichtert auch die Verifikation und Falsifikation von Implikationen in einem Modell. Daß z.B. das Identitätsschema

Id. $$A \to A$$

gilt, ist jetzt eine triviale Einsicht. Die Schemata (74) und (75) sind in A-Rahmen jedoch nicht gültig.

- Die Formel $P \to (Q \to P)$ falsifizieren wir wie folgt. Wir betrachten einen Rahmen mit der Punktmenge $\{0, a, b, c, d\}$; R erfülle die Bedingungen $R0ab$ (also $a \leq b$) und $Rbcd$. (Für (R1) fügen wir $(\forall x)R0xx$, für (R2) $Racd$ hinzu. Der Leser überzeuge sich davon, daß dann auch (R3-4) erfüllt sind.) Nun interpretieren wir P und Q auf diesem Rahmen so: $I(P, a) = 1$, $I(Q, c) = 1$, $I(P, d) = 0$. (Die Wahrheit von P wird von a nach b weitergegeben, aber das muß uns nicht weiter interessieren.) Dann wird $Q \to P$ am Punkt b falsch. (Denn es gibt Punkte x und y – nämlich c und d – mit $Rbxy$, $x \models Q$ und $y \not\models P$.) Andererseits ist P am Punkte a wahr. Als kann die Ausgangsformel in diesem Modell am Punkt 0 nicht wahr sein.
- Im selben Rahmen können wir auch $P \to (Q \to Q)$ falsifizieren. Nur betrachten wir jetzt eine Interpretation mit $I(P, a) = 1$ und $I(Q, c) = 1$ (wie zuvor), sowie $I(Q, d) = 0$. Dann wird $Q \to Q$ am Punkt b falsch. Punkt b ist also ein nicht-normaler Punkt, an dem auch logisch gültige Formeln falsch werden können (vgl. p. 358).

Wie nützlich das Korollar zum Lemma 9 ist, sehen wir gleichfalls im Beweis des nächsten Lemmas, in dem wir schließlich auch Gebrauch von (R3-4) machen.

LEMMA 11. *Die folgenden Schemata sind gültig bzw. übertragen die Gültigkeit der Prämissen auf die Konklusion:*

B. $$(A \to B) \to ((C \to A) \to (C \to B))$$

CB. $$(A \to B) \to ((B \to C) \to (A \to C))$$

MP. $$\frac{A \quad A \to B}{B}$$

Beweis. *Ad* B. Wir nehmen an,

$$(1) \qquad a \models A \to B$$

und zeigen nach dem Korollar, daß $a \models (C \to A) \to (C \to B)$. Also nehmen wir ferner an, daß

$$(2,3) \qquad Rabc \quad \text{und} \quad b \models C \to A.$$

(Nun zu zeigen: $c \models C \to B$.) Nach Annahme von

$$(4,5) \qquad Rcde \quad \text{und} \quad d \models C,$$

genügt es schließlich zu zeigen, daß $e \models B$. — (3) bedeutet

$$(6) \qquad \forall xy : Rbxy \;\&\; x \models C \;\Rightarrow\; y \models A.$$

Aus (2) und (4) folgt

$$(7) \qquad R^2(ab)de;$$

also nach (R3),

$$(8) \qquad R^2a(bd)e, \text{ d.h. } \exists y : Rbdy \;\&\; Raye.$$

Aus (5), (6) und (8) folgt dann

$$(9) \qquad \exists y : Raye \;\&\; y \models A.$$

Nun bedeutet (1),

$$(10) \qquad \forall yz : Rayz \;\&\; y \models A \;\Rightarrow\; z \models B.$$

So folgt aus (9) und (10), $e \models B$, wie gewünscht.

Ad CB. Übung! Man benutze (R3*).

Ad MP. Offensichtlich. ■

In der Kripke-Semantik der Modallogik haben wir die Frage gestellt, welche logische Theorie die "reinen" Rahmen beschreibt, d.h. diejenigen Rahmen, die unter keinerlei strukturellen Bedingungen für die Zugangsrelation stehen. Natürlich können wir diese Frage auch im Falle der Routley-Meyer Semantik stellen – und die Antwort ist gewissermaßen enttäuschend: Die Theorie der "reinen" A-Rahmen ist durch die leere Menge gegeben.

Schon für die Gültigkeit von $P \to P$ wird atomare Weitergabe benötigt. Auch die Theorie des Folgerns ist in solchen Rahmen trivial. Zum Beispiel ist der atomare Modus Ponens – $P, P \to Q/Q$ – ohne atomare Weitergabe und die Bedingung $R000$ nicht zu haben.

Wir benötigen also Bedingungen für die Relation R, damit die Interpretationsklausel für $\to$ überhaupt eine Logik erzeugen kann. Aber warum dürfen die in der Definition angegebenen Bedingungen als minimale Bedingungen gelten? Nun, das Zeichen $\to$ soll eine Art logisches "Wenn ..., dann ..." ausdrücken. Dafür müssen zwei Voraussetzungen erfüllt sein:

· Unter dieser beabsichtigten Lesart kann $A \to B$ in einem Modell nicht wahr sein, wenn es darin Punkte gibt, an denen A wahr aber B falsch sind. Solche Punkte würden $A \to B$ im Modell falsifizieren.

· Wenn $A \to B$, dann sollte die Wahrheit von A *hinreichend* für die Wahrheit von B sein. Genau diese Bedingung wird durch die Schnitteigenschaft ausgedrückt: Wenn $A \to B$ wahr ist an einem Punkt, dann sollten auch $(B \to C) \to (A \to C)$ und $(C \to A) \to (C \to B)$ an diesem Punkt wahr sein. Ist das nicht der Fall – wie bei kontrafaktischen Konditionalen –, dann ist die naheliegendste Erklärung, daß A nicht "unter allen Umständen" hinreichend für B ist.

Nun haben wir im Korollar zu Lemma 9 gesehen, daß (R1-2) zusammen mit der Weitergabebedingung gewährleistet, daß die erste Voraussetzung erfüllt ist. Die in der zweiten Voraussetzung genannte Schnitteigenschaft ist unschwer in den Schemata B und CB zu erkennen. Diese sind aber mit den Bedingungen (R3) und (R3*) gekoppelt: B gilt in einem Rahmen genau dann, wenn der Rahmen die Bedingung (R3) erfüllt; ditto CB vis-à-vis (R3*).[18] Die Bedingungen für A-Rahmen sind also notwendig, damit in den darauf aufliegenden Modellen $\to$-Formeln als Implikationen interpretiert werden können.

[18] Es gibt zwei Strategien, in der allgemeinen Definition eines A-Rahmens ohne (R3) und (R3*) auszukommen. Erste Strategie: Ersetze in der Definition eines Rahmens den Nullpunkt durch eine Menge sogenannter normaler Punkte und richte den Rest entsprechend ein. Es entsteht so die Definition eines *unreduzierten* Rahmens. (Unsere Rahmen in Definiton 6) sind, in dieser Terminologie, reduziert.) Die Theorie unreduzierter Rahmen ist wie $\mathbf{C}_\to$ unten, nur daß die Schemata B und CB zu Regeln abgeschwächt sind. Das ist die implikative Basislogik $\mathbf{B}_\to$ – der Namenskonvention in [254] folgend – für unreduzierte Rahmen. Unreduzierte Rahmen werden z.B. in [210, 227] und besonders detailliert in [254] dargestellt. Nachteil: Implikation an einem normalen Punkt bedeutet nicht mehr, daß alle Antezedens-Welten auch Konsequens-Welten sind. — Zweite Strategie: Bleibe bei reduzierten Rahmen und ergänze das Vollständigkeitsargument durch eine Metabewertung wie in Slaney [268]. Nachteil: Keiner – nur sprengen Metabewertungen den Rahmen unserer Darstellung. Unter jeder Strategie tauchen (R3) und (R3*) natürlich wieder auf, sobald die Logik um B und CB erweitert wird – und das geschieht in der Regel schon im ersten Erweiterungsschritt.

Unter der Logik $\mathbf{C}_\rightarrow$ wollen wir die Menge der Theoreme verstehen, die sich aus den folgenden Schemata ableiten lassen: [19]

Das System $\mathbf{C}_\rightarrow$

Id.	$A \rightarrow A$
B.	$(A \rightarrow B) \rightarrow ((C \rightarrow A) \rightarrow (C \rightarrow B))$
CB.	$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$
MP.	$\dfrac{A \quad A \rightarrow B}{B}$

Wir wollen jetzt beweisen, daß $\mathbf{C}_\rightarrow$ genau die Theorie logischer Implikationen in der Klasse $\mathbb{A}$ aller A-Rahmen ist. D.h. alle Theoreme von $\mathbf{C}_\rightarrow$ sind in $\mathbb{A}$ gültig; und jede in $\mathbb{A}$ gültige Formel ist ein Theorem von $\mathbf{C}_\rightarrow$. Eine Hälfte dieser Behauptung folgt unmittelbar aus Lemma 11:

Satz 12. (Richtigkeit) *Jedes Theorem von* $\mathbf{C}_\rightarrow$ *ist in der Klasse* $\mathbb{A}$ *aller Rahmen gültig.*

Es bleibt der Nachweis der Vollständigkeit von $\mathbf{C}_\rightarrow$ im Hinblick auf A-Rahmen. Wir beginnen mit einer Definition. Wenn S eine Menge von Sätzen ist, dann wollen wir eine Formelmenge X *S-implikativ* nennen, wenn sie unter Implikation in S abgeschlossen ist, d.h. wenn gilt:

$$A \in X \;\&\; A \rightarrow B \in S \;\Rightarrow\; B \in X.$$

Unter einem *kanonischen Modell* für $\mathbf{C}_\rightarrow$ wollen wir eine Struktur

$$(W^{\mathbf{C}_\rightarrow}, 0^{\mathbf{C}_\rightarrow}, R^{\mathbf{C}_\rightarrow}, I^{\mathbf{C}_\rightarrow})$$

verstehen, die wie folgt bestimmt sein soll:

> (Das hochgestellte "$\mathbf{C}_\rightarrow$" lassen wir jetzt fort, ohne dabei zu vergessen, daß erst noch zu zeigen ist, daß kanonische Modelle tatsächlich Modelle auf A-Rahmen im Sinne der Definition 6 sind!)

1. W ist eine nichtleere Menge von 0-implikativen Formelmengen;

[19] $\mathbf{C}_\rightarrow$ hat nichts mit den $\mathbf{C}$-Systemen von da Costa im Kapitel über parakonsistente Logik zu tun. Wir leiten die Bezeichnung hier aus [254] ab.

2. 0 ist eine 0-implikative Menge, die alle Theoreme von $\mathbf{C}_{\rightarrow}$ enthält;[20]
3. $R \subseteq W^3$ so, daß $Rabc$ gdw für alle Formeln A und B: $A \rightarrow B \in a \;\&\; A \in b \;\Rightarrow\; B \in c$;
4. $I : \mathrm{ATM} \times W \longrightarrow \{0,1\}$ so, daß $I(P,a) = 1$ gdw $P \in a$.
   Ferner definieren wir $\models^{\mathbf{C}_{\rightarrow}}$ so: $a \models^{\mathbf{C}_{\rightarrow}} A$ gdw $A \in a$.

Aus der Definition geht unmittelbar hervor, daß es für jedes Nicht-Theorem F von $\mathbf{C}_{\rightarrow}$ ein kanonisches Modell gibt mit $F \notin 0$. Erfüllen R und I in einem solchen Modell die R-Bedingungen sowie die Bedingung (Atm $\leq$) der Modell-Definition 6? Und erfüllt die Relation $\models$ schließlich die Interpretationsbedingung ($\rightarrow$)? Wenn es uns gelingt, dies zu zeigen, dann handelt es sich bei kanonischen Modellen tatsächlich um Modelle im Sinne unserer Definition (Hauptlemma). Aber dann haben wir folgendes Resultat: Wenn F in $\mathbf{C}_{\rightarrow}$ nicht beweisbar ist, dann gibt es ein $\mathbb{A}$-Modell, in dem F nicht wahr ist – also ist F in der Rahmenklasse $\mathbb{A}$ nicht gültig. Das ist die gewünschte Umkehrung des Satzes 12. Im jetzt verhandelten Fall einer rein implikativen Sprache ist der Beweis besonders einfach.

LEMMA 13. (Hauptlemma) *Jedes kanonische Modell für $\mathbf{C}_{\rightarrow}$ erfüllt die Bedingungen für Modelle auf A-Rahmen (Definition 6).*

BEWEIS. *I. Die Rahmenbedingungen.*

Ad Reflexivität (R1) und Weitergabe (Atm $\leq$): An dieser Stelle ist die folgende Beobachtung über kanonische Modelle hilfreich:

$$a \leq b \text{ gdw } a \subseteq b.$$

Von rechts nach links ist das trivial. In der anderen Richtung nehmen wir an, daß $A \in a$. Da $A \rightarrow A \in \mathbf{C}_{\rightarrow} \subseteq 0$ und a 0-implikativ ist, so ist $A \in b$. Reflexivität und (allgemeine) Weitergabe sind jetzt offensichtlich erfüllt.

Ad Monotonie (R2): Wir nehmen an (1) $a \leq b$, also $a \subseteq b$, und (2) $Rbcd$ im kanonischen Modell. Zu zeigen ist $Racd$. So nehmen wir denn ferner an (3) $A \rightarrow B \in a$ sowie (4) $A \in c$. (Jetzt ist zu zeigen, daß $B \in d$.) Aus (1) und (3) folgt (5) $A \rightarrow B \in b$. Aus (2) und (5) folgt (6) $A \in c \Rightarrow B \in d$. Aus (4) und (6) folgt dann das Gewünschte.

[20] Um die Vollständigkeit von $\mathbf{C}_{\rightarrow}$ nachzuweisen, könnten wir an dieser Stelle einfach $0 = \mathbf{C}_{\rightarrow}$ setzen. Aber für bestimmte Erweiterungen von $\mathbf{C}_{\rightarrow}$ wird das nicht mehr möglich sein. Deshalb wählen wir schon hier die schwächere Bedingung $\mathbf{C}_{\rightarrow} \subseteq 0$.

Ad Assoziativität (R3): Wir nehmen an, daß (1) $\exists x Rabx$ & $Rxcd$ (und zeigen, daß $\exists y Rbcy$ & $Rayd$). Aus (1) folgt, daß es ein x gibt so, daß $\forall ABCD$:

$$A \to B \in a \;\&\; A \in b \;\Rightarrow\; B \in x; \tag{2}$$

$$C \to D \in x \;\&\; C \in c \;\Rightarrow\; D \in d. \tag{3}$$

Es sei nun die Menge y so definiert:

(Dy) $\qquad F \in y$ gdw $\exists B \in b$, $\exists C \in c:\; B \to (C \to F) \in 0$.

1. y ist 0-implikativ. Denn, angenommen (4) $F \to D \in 0$ und (5) $F \in y$. Letzteres bedeutet nach (Dy), daß es B und C gibt mit (6) $B \in b$, (7) $C \in c$ und (8) $B \to (C \to F) \in 0$. Da b 0-implikativ ist, folgt aus (6) und (8), (9) $C \to F \in b$. Nun ist aber aufgrund von CB (10) $(F \to D) \to ((C \to F) \to (C \to D)) \in 0$. Aus (4) und (10) folgt, (11) $(C \to F) \to (C \to D) \in 0$. Aus (9) und (11) folgt, (12) $C \to D \in b$. Ferner haben wir (13) $(C \to D) \to (C \to D) \in 0$. Also gibt es Formeln B' $(= C \to D)$ und C mit $B' \in b$ (nach (12)), $C \in c$ (nach (7)) und $B' \to (C \to D)$ (nach (13)), d.h. $D \in y$.
2. *Rbcy*. Denn, angenommen (4) $A \to F \in b$ und (5) $A \in c$. (Zu zeigen: $F \in y$.) Da (6) $(A \to F) \to (A \to F) \in 0$ und b 0-implikativ ist, so folgt aus (4), daß es Formeln B und C gibt (nämlich $B = (A \to F)$ und $C = A$) so, daß $B \in b$, $C \in c$ und $B \to (C \to F) \in 0$. Aber dann ist nach (Dy), $F \in y$, wie gewünscht.
3. *Rayd*. Denn, angenommen (4) $F \to A \in a$ und (5) $F \in y$. (Zu zeigen: $A \in d$.) Nach (Dy) folgt aus (5), (6) $\exists BC : B \in b \;\&\; C \in c \;\&\; B \to (C \to F) \in 0$. Aufgrund von CB haben wir (7) $(F \to A) \to ((C \to F) \to (C \to A)) \in 0$. Da a 0-implikativ ist, folgt aus (4) und (7), (8) $(C \to F) \to (C \to A) \in a$. Ebenso folgt aus (6), (9) $C \to F \in b$. Somit aus (8) und (9) nach (2): (10) $C \to A \in x$. Aus (10) und (6) folgt schließlich nach (3) das gewünschte $A \in d$.

Ad Assoziativität* (R3*): Ähnlich wie (R3), mit entsprechend permutierter Definition von y und Einsatz von B statt CB.

II. Die Modellbedingungen.

Ad $(\to)$. Zu zeigen ist:

$$\begin{aligned} A \to B \in a \text{ gdw } & \forall bc : [Rabc] \;\&\; A \in b \;\Rightarrow\; B \in c, \\ \text{gdw } & \forall bc : [\forall CD : C \to D \in a \;\&\; C \in b \;\Rightarrow\; D \in c] \\ & \qquad \&\; A \in b \;\Rightarrow\; B \in c. \end{aligned}$$

Von links nach rechts ist das offensichtlich (A für C, B für D). Für die umgekehrte Richtung buchstabieren wir zunächst um und kontraponieren. Wir möchten jetzt zeigen, daß

$$A \to D \notin a \Rightarrow \exists bc : Rabc \;\&\; A \in b \;\&\; D \notin c.$$

Wir definieren:

(Db) $\qquad b := \{B : \; A \to B \in 0\};$

(Dc) $\qquad c := \{C : (\exists B \in b)\, B \to C \in a\}.$

1. *b ist 0-implikativ.* Angenommen (1) $B_1 \in b$ und (2) $B_1 \to B_2 \in 0$. Dann folgt aus (1) nach (Db) (3) $A \to B_1 \in 0$. Nun gilt aufgrund von CB,

   (4) $\qquad A \to B_1 \in 0 \;\&\; B_1 \to B_2 \in 0 \Rightarrow A \to B_2 \in 0.$

   So folgt aus (2-4), (5) $A \to B_2 \in 0$, d.h. nach (Db), $B_2 \in b$, wie gewünscht.

2. *c ist 0-implikativ.* Angenommen (1) $C_1 \in c$ und (2) $C_1 \to C_2 \in 0$. Aus (1) folgt nach (Dc), daß (3) $B \to C_1 \in a$, für ein $B \in b$. Nun gilt aufgrund von B,

   (4) $\qquad C_1 \to C_2 \in 0 \;\&\; B \to C_1 \in a \Rightarrow B \to C_2 \in a.$

   So folgt aus (2-4), (5) $B \to C_2 \in a$ mit $B \in b$. Also nach (Dc), $C_2 \in c$.

3. *$Rabc$.* Aus den Annahmen $B \to C \in a$ und $B \in b$ folgt unmittelbar nach (Dc), daß gewünschte $C \in c$.

4. $A \in b$. Aufgrund von I und (Db).

5. $D \notin c$. Angenommen – für *reductio* – (1) $D \in c$. Dann gibt es nach (Dc) ein $B \in b$ mit (2) $B \to D \in a$. Da $B \in b$, ist nach (Db), (3) $A \to B \in 0$. Daraus folgt mit B, (4) $(B \to D) \to (A \to D) \in 0$. Daher aus (2), (5) $A \to D \in a$ – was unserer Voraussetzung widerspricht. ∎

Mit dem Beweis dieses Lemmas ist unser Nachweis der Vollständigkeit von $\mathbf{C}_{\to}$ im Hinblick auf die Menge der Formeln, die in Modellen für $\to$ auf A-Rahmen gültig sind, beendet. Wir halten das Resultat in einem Satz fest.

Satz 14. *Jede Formel, die in (Modellen für $\to$ auf Rahmen aus) der Klasse $\mathbb{A}$ aller Rahmen gültig ist, ist ein Theorem von $\mathbf{C}_\to$.*

Wir wollen $Th_\to(\mathbb{A})$, die Menge der in Modellen für $\to$ auf Rahmen in $\mathbb{A}$ gültigen Formeln, die *implikative Theorie* von $\mathbb{A}$ nennen. Die Sätze 12 und 14 besagen dann, daß das System $\mathbf{C}_\to$ die implikative Theorie von $\mathbb{A}$ angibt:

$$\mathbf{C}_\to = Th_\to(\mathbb{A}).$$

In der Terminologie, die wir im Kapitel über Modallogik verwendet haben: Die Menge aller A-Rahmen *determiniert* die Logik $\mathbf{C}_\to$.

Die Methode der kanonischen Modelle liefert nicht immer auf so einfache Weise einen Vollständigkeitsbeweis. Im Falle von $\mathbf{C}_\to$ sind die Dinge so einfach, weil wir nicht die Interpretationsbedingungen für die anderen Junktoren, insbesondere für die Negation und Disjunktion, in kanonischen Modellen verifizieren müssen. Die Punkte in einem kanonischen Modell müssen nur unter Implikation am Nullpunkt abgeschlossen sein, und den Nullpunkt des Modells können wir in diesem Fall sogar mit der Theoremmenge von $\mathbf{C}_\to$ gleichsetzen. Wenn wir später zu einer reicheren Sprache übergehen, können wir diese Einfachheit nicht mehr erwarten. Die Konstruktion geeigneter Punkte wird aufwendiger. Dennoch bleibt die Struktur des Argumentes gleich. Der einfache Beweis für die Basislogik $\mathbf{C}_\to$ stellt daher besonders gut die Behandlung der Implikation in den A-Rahmen heraus – an dieser ändert sich im wesentlichen nichts, wenn wir die Sprache erweitern.

5. Mehr Implikationen

Ausgehend vom Basissystem $\mathbf{C}_\to$ können wir jetzt den Effekt von Einschränkungen der Rahmen betrachten. Darunter wollen wir hier Bedingungen verstehen, unter welche die Relation R gestellt werden kann. Wird die Rahmenklasse – innerhalb derer für die Bestimmung der gültigen Formeln die Modelle variiert werden – kleiner, so steigt gewissermaßen die Chance einer Formel gültig zu sein. Neue Bedingungen auf Rahmen können so Formeln gültig machen, die es zuvor nicht waren. Uns interessieren in diesem Zusammenhang insbesondere Paare von Rahmenbedingungen und Formelschemata, die, in einem gleich zu definierenden Sinne, einander "entsprechen".

Die Bedeutung des System $\mathbf{C}_\to$ erschöpft sich darin, daß es die Klasse der A-Rahmen charakterisiert. Die A-Rahmen bilden einen natürlichen Ausgangspunkt für die Interpretation des Junktors $\to$ im Sinne einer logischen

Implikation. Viele Prinzipien der Implikation, die der Idee einer Relevanzlogik eigentlich nicht widersprechen, sind in $\mathbf{C}_{\rightarrow}$ jedoch nicht ableitbar. (Das schließt natürlich nicht aus, daß solche weiteren Prinzipien aus anderer Perspektive unter Verdacht geraten können.) In der nächsten Tafel sehen wir eine Auswahl von Implikations-Schemata, die keine Theoreme von $\mathbf{C}_{\rightarrow}$ sind.

Schemata

W (Kontraktion).	$(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$
C (Permutation).	$(A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$
CI (Ablösung).	$A \rightarrow ((A \rightarrow B) \rightarrow B)$
M (Mingle).	$A \rightarrow (A \rightarrow A)$
K (Abschwächung).	$A \rightarrow (B \rightarrow A)$

Man beachte, daß diese Schemata nicht unabhängig voneinander sind! Tatsächlich sind C und CI äquivalent, gegeben Id, B und CB. Siehe dazu die Anmerkung auf p. 375.

> *Übung*: Der Leser möge versuchen, diese Schemata in $\mathbb{A}$-Modellen zu falsifizieren. Auf diese Weise wird er selbst auf die gleich zu nennenden Rahmen-Bedingungen kommen, die den Schemata entsprechen.

Für jedes dieser Schemata wollen wir eine Rahmenbedingung finden so, daß das Schema genau in solchen Rahmen gültig ist, welche die Bedingung erfüllen. Eine Korrespondenz in diesem Sinne zwischen Schema und Bedingung ist aber für sich nicht hinreichend, um die Vollständigkeit bestimmter Erweiterungen von $\mathbf{C}_{\rightarrow}$ bezüglich bestimmter Rahmenklassen anzuzeigen. Dazu müssen wir etwas Stärkeres beweisen:

(a) Das Schema S ist in allen Rahmen gültig, welche die Bedingung erfüllen; und

(b) wenn eine Logik das Schema S enthält, dann erfüllt das kanonische Modell für diese Logik die Bedingung.

Diese Relation zwischen einem Schema und einer Rahmenbedingung haben wir im Kapitel über Modallogik *kanonische Korrespondenz* genannt. (Man erinnere sich an die Diskussion dort im Anschluß an die Definition 19.) Die Klausel (a) behauptet die Richtigkeit des Schemas bezüglich von Rahmen, die unter der Bedingung stehen. Klausel (b) beschreibt genau das, was für ein Henkin'sches Vollständigkeitsargument benötigt wird; vgl. das Hauptlemma 13. Da uns hier die Beziehung zwischen Schema und Bedingung nur im Hinblick auf Richtigkeits- und Vollständigkeitsnachweise bestimmter Logiken bzw. Rahmenklassen interessiert, sei mit *Korrespondenz* im weiteren *kanonische* Korrespondenz gemeint.

Zur weiteren Verständigung wollen wir jetzt ein wenig Notation vereinbaren (bzw. daran erinnern):

- Wenn (Bed) eine Bedingung auf $\mathbb{A}$-Rahmen ist, dann sei $Bed(\mathbb{A})$ die Klasse der A-Rahmen, welche die Bedingung erfüllen.
- Wenn A ein Formelschema ist, dann sei $Rm(A)$ die Klasse der $\mathbb{A}$-Rahmen, in denen A gültig ist.
- Wenn $\mathbb{R}$ eine Rahmenklasse ist, dann sei $Th_{\rightarrow}(\mathbb{R})$ die implikative Theorie von $\mathbb{R}$.
- Wenn $\mathbf{L}$ ein logisches Axiomensystem ist, dann sei $\mathbf{L}+A$ die Erweiterung von $\mathbf{L}$ um das Axiomenschema A.

Definition 15. (Determination) Das Schema A *determiniert* genau dann die Rahmenbedingung (Bed), wenn

(Det.) $$Th_{\rightarrow}(Bed(\mathbb{A})) = \mathbf{C}_{\rightarrow} + A.$$

Die Gleichung (Det.) besagt natürlich nichts anderes als, daß die Rahmenklasse $Bed(\mathbb{A})$ die Logik $\mathbf{C}_{\rightarrow} + A$ determiniert, d.h. letztere richtig und vollständig bezüglich $Bed(\mathbb{A})$ ist. Wie soeben erklärt, ist daher (Det.) äquivalent zur Behauptung, daß (Bed) und A miteinander (kanonisch) korrespondieren.

Es folgt die Liste der Rahmenbedingungen, die mit den oben genannten Schemata korrespondieren.

Rahmenbedingungen

(c) $$R^2(ab)cd \Rightarrow R^2(ac)bd$$

(ci) $$Rabc \Rightarrow Rbac$$

(w) $$Rabc \Rightarrow R^2(ab)bc$$

(m) $$Rabc \Rightarrow (a \leq c \text{ oder } b \leq c)$$

(k) $$Rabc \Rightarrow a \leq c$$

Lemma 16. (Korrespondenz) *Das Schema S (aus der Liste der Schemata) korrespondiert mit der Bedingung (s) (aus der Liste der Rahmenbedingungen).*

Beweis. Wir illustrieren den Beweis anhand zweier einfacher Paare: (CI, (ci)) und (K,(k)).

CI korr. (ci). Wir betrachten ein beliebiges Modell unter der Bedingung (c) und verfizieren das Schema CI wie folgt: Wir nehmen an, daß (1) $a \models A$. Da zu zeigen ist, daß $a \models (A \to B) \to B$, so nehmen wir ferner an, daß (2) $Rabc$ und (3) $b \models A \to B$. Aus (2) folgt nach (ci), (4) $Rbac$, was im Verein mit (1) und (3) das gewünschte $c \models B$ ergibt.

In umgekehrter Richtung zeigen wir, daß jedes kanonische Modell für $\mathbf{C}_\to$+CI die Bedingung (ci) erfüllt. Wir nehmen also an, daß (1) $Rabc$ und zeigen, daß $Rbac$. D.h. unter der Annahme (2) $A \to B \in b$ und (3) $A \in a$ ist jetzt zu zeigen, daß $B \in c$. Wir wissen, daß (4) $A \to ((A \to B) \to B) \in 0$, da jetzt $\mathbf{C}_\to + \mathrm{CI} \subseteq 0$. Aber dann aus (3) und (4), (5) $(A \to B) \to B \in a$. Und aus (1), (2) und (5) folgt $B \in c$, wie gewünscht.

K korr. (k). Angenommen (1) $a \models A$; zu zeigen ist $a \models B \to A$. Wir nehmen also ferner an (2) $Rabc$ (und zeigen jetzt $c \models A$; die ebenfalls zur Verfügung stehende Annahme $b \models B$ ist "irrelevant"). Aus (2) folgt nach (k), (3) $a \leq c$. Dann, nach Lemma 9, $c \models A$, wie gewünscht.

Für die umgekehrte Richtung nehmen wir im kanonischen Modell an, daß (1) $Rabc$ und zeigen, daß $a \leq c$, d.h. im kanonischen Modell $a \subseteq c$. So nehmen wir ferner an, (2) $A \in a$ (und zeigen, daß $A \in c$). Aus (1) folgt, (3) $B \to A \in a$ & $B \in b \Rightarrow A \in c$. Aber da a unter Implikation in $0 \subseteq \mathbf{C}_\to$+K abgeschlossen ist, folgt aus (2), (4) $B \to A \in a$ und so aus (3) und (4), (5) $B \to b \Rightarrow A \in c$, für beliebiges B. Da $b \neq \emptyset$ erhalten wir so $A \in c$, wie gewünscht.

Etwas schwieriger zu verifizieren sind Bedingungen, in denen im Konsequens ein Existenzquantor vorkommt. Der Fall (C,(c)) verfährt nach dem Muster von (B,(R3)), was wir im Beweis des Hauptlemmas für $\mathbf{C}_\to$ bereits vorgeführt haben. Das Paar (W,(w)) ist ebenfalls etwas vertrackt. Wir wollen uns hier mit der Beobachtung begnügen, daß (w) aus den Bedingungen (R3) und (ci) sowie der Idempotenz-Bedingung $Raaa$ (die wir erst im Zusammenhang mit dem System **C** betrachten werden) folgt:

Angenommen $Rabc$. Dann $Rbac$ nach (ci). Es folgt im Verein mit $Rbbb$, daß $R^2(bb)ac$. Daraus folgt nach (R3): $R^2b(ba)c$, d.h. $\exists x : Rbax$ & $Rbxc$. Wir kommutieren nach (ci) in beiden Konjunkten und erhalten so $\exists x : Rabx$ & $Rxbc$, d.h. $R^2(ab)bc$, wie gewünscht. ∎

Aus den Korrespondezen des Lemmas lassen sich nun Vollständigkeitsresultate für Erweiterungen von $\mathbf{C}_\to$ nach einem Baukastenprinzip ableiten. Es sei z.B. $\mathbf{C}_\to$+(C,W) die Logik, die aus $\mathbf{C}_\to$ durch Erweiterung um die Axiomenschemata C und W ensteht. Dann ist die Menge der Theoreme dieser Logik genau die Theorie der Rahmen, welche die Bedingungen (c) und (w) erfüllen. Bei einer Erweiterung von $\mathbf{C}_\to$ um C wird übrigens eines der Axiome B oder CB redundant. Denn – wie es ja schon der Name andeutet – CB entsteht offensichtlich aus B durch die Permutation C der Antezedensformeln $A \to B$ und $C \to A$, und sinngemäß auch umgekehrt. Entsprechend erfüllt jeder Rahmen unter den Bedingungen (c) und (R3) (bzw. (R3*)) auch die Bedingung (R3*) (bzw. (R3)). Allgemein haben wir den folgenden Satz:

SATZ 17. *Es seien $S_1, ..., S_n$ Schemata aus der Liste und $(s_1), ..., (s_n)$ korrespondierende Rahmenbedingungen (Lemma 16!). Dann*

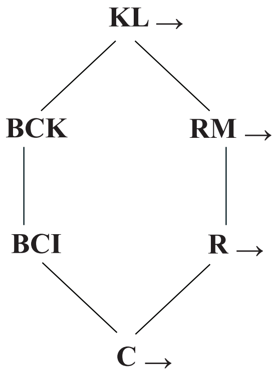
$$\mathbf{C}_{\to} + (S_1, ..., S_n) = Th(\mathbb{A}|s_1, ..., s_n),$$

wobei $\mathbb{A}|s_1, ..., s_n$ die Klasse der A-Rahmen ist, welche die Bedingungen $(s_1), ..., (s_n)$ erfüllen.

Von den Logiken, die von dem Satz erfaßt werden, seien hier nur die folgenden erwähnt:

$$\begin{array}{ll}
\mathbf{R}_{\to} & = \mathbf{C}_{\to}\text{+(C,W)} \\
\mathbf{RM}_{\to} & = \mathbf{C}_{\to}\text{+(C,W,M)} \\
\mathbf{BCI} & = \mathbf{C}_{\to}\text{+C} \\
\mathbf{BCK} & = \mathbf{C}_{\to}\text{+(C,K)} \\
\mathbf{KL}_{\to} & = \mathbf{C}_{\to}\text{+(C,W,K)}
\end{array}$$

Alle diese Logiken, außer **BCI** und **BCK**, sind implikative Fragmente von bekannten Logiken in $\{\neg, \wedge, \vee, \to\}$; **BCI** und **BCK** werden gewöhnlich als rein implikative Logiken aufgefaßt. **R** ist die Relevanzlogik, auf die wir in diesem Kapitel zusteuern. **RM** ist eine Erweiterung von **R** um das sogenannte Mingle-Axiom M. **RM** hat eine besonders einfache Modelltheorie, erfüllt aber nicht die Überschneidungsbedingung aufgrund von Theoremen wie $A \wedge \neg A \to B \vee \neg B$. In [10] wird **RM** ausgiebig untersucht. **BCK** ist eine Abschwächung der klassischen Logik **KL** um das Kontraktionsschema W, und **BCI** eine weitere Abschwächung. Diese Logiken sind aus verschiedenen Perspektiven, die uns jedoch vom Gegenstand der Relevanzlogik ablenken würden, interessant; siehe z.B. [143, Kap. 7.25]. Die Inklusionsverhältnisse zwischen diesen Logiken gibt das folgende Diagramm wieder. (Wenn ein Pfad von unten nach oben eine Logik mit einer anderen verbindet, dann ist jene in dieser echt enthalten.)



Anmerkung. All diese Logiken können wir auch axiomatisieren, indem wir das Permutations-Schema C durch das Ablösungsschema CI ersetzen. Denn im Basissystem $\mathbf{C}_\to$ haben wir die Schemata Id, B und CB zur Verfügung, mit deren Hilfe C und CI wechselseitig ableitbar sind:

CI $\Rightarrow$ C:

(1)	$B \to ((B \to C) \to C)$	CI
(2)	$(((B \to C) \to C) \to (A \to C)) \ \to \ (B \to (A \to C))$	1, B, MP
(3)	$(A \to (B \to C)) \to (((B \to C) \to C) \to (A \to C))$	CB
(4)	$(A \to (B \to C)) \to (B \to (A \to C))$	2,3, Trans.

C $\Rightarrow$ CI folgt aus der Id-Instanz $(A \to B) \to (A \to B)$ durch Permutation C der Pämissen. — Aufgrund der Korrespondenzen können wir daher schließen, daß in A-Rahmen die Bedingungen (c) und (ci) äquivalent sind. Die direkte Verifikation sei dem Leser zur Übung überlassen.

Die allgemeine Form der Interpretation von Konditionalen. Wir kommen nun auf die auf p. 360 gestellte Frage zurück, warum die dreistellige Relation R ein geeignetes Mittel ist, die Bedeutung eines konditionalen Satzes zu modellieren. Dazu erinnern wir uns an die Semantik der Konditionalsätze, die wir im vierten Kapitel im Abschnitt über variabel relationale Rahmen vorgestellt haben. Die Grundidee war recht einfach: Ein (modales) Konditional $A \to B$ behauptet die Wahrheit des Konsequens B mit einer gewissen Notwendigkeit. Diese Notwendigkeit muß irgendwie durch das Antezedens A konditioniert sein. Man könnte auch sagen: Ein Konditional, das mit Notwendigkeit wahr sein soll, behauptet die konditionale Notwendigkeit des Konsequens – wobei das Antezedens die Kondition benennt. Wenn wir nun kategorische Notwendigkeit mit einer *zweistelligen* Relation modellieren, dann brauchen wir für konditionale Notwendigkeit offenbar eine *dreistellige* Relation, in welcher der semantische Wert des Antezedens den dritten Parameter abgibt. Sei nun S ein solche dreistellige Relation, d.h. zwei Stellen werden von Welten eingenommen – wie in einer Zugangsrelation in der Modallogik – und eine Stelle wird von einer Punktemenge eingenommen; das werden, wie wir gleich festlegen, die Punkte sein, an denen das Antezedens des zu beurteilenden Konditionals wahr ist. Es wird sich als sinnvoll erweisen, die Koordinaten so zu reihen:

$$S \subseteq W \times \wp(W) \times W.$$

Dann ist die folgende Äquivalenz so etwas wie eine Blaupause für die Interpretation jeder Art (modaler) Konditionalsätze:[21]

$$(S \rightarrow) \qquad a \models A \rightarrow B \text{ gdw } \forall z : Sa[\![A]\!]z \Rightarrow z \models B.$$

(Tatsächlich fallen auch materiale, d.h. nicht-modale Konditionale unter dieses Schema. Dieser Fall entsteht, wenn die Relation S durch die Bedingung $SaXb \Rightarrow a = b$ trivialisiert wird.) Wir zeigen nun, daß auch die Interpretation von Konditionalen in A-Rahmen mit der Bedingung

$$(R \rightarrow) \qquad a \models A \rightarrow B \text{ gdw } \forall yz : Rayz \;\&\; y \models A \Rightarrow z \models B$$

unter dieses Schema fällt.

Wir betrachten ein Modell $\mathcal{M}^R = (W, 0, R, I)$ und definieren darin eine Relation $S \subseteq W \times \wp(W) \times W$ so:

$$(\text{Def.}\, S) \qquad SaYb \text{ gdw } \exists y : y \in Y \;\&\; Rayc.$$

Dann können wir zeigen, daß die jeweils rechten Seiten von $(S \rightarrow)$ und $(R \rightarrow)$ in $\mathcal{M}^R$ äquivalent sind.

BEOBACHTUNG 18. *In $\mathcal{M}^R$ gilt für alle $a \in W$ und alle $Y, Z \subseteq W$: $(\forall z : SaYz \Rightarrow z \in Z)$ gdw $(\forall yz : Rayz \;\&\; y \in Y \Rightarrow z \in Z)$.*

BEWEIS. ($\overline{R}abc$ sei kurz für $(a, b, c) \notin R$.)

LR: Wir nehmen an (1) die linke Seite, sowie (2) $Rabc$ und (3) $b \in Y$. Ferner – für *reductio* – (4) $c \notin Z$. Aus (1) und (4) folgt (5) $\overline{S}aYc$, d.h. (Def. S) (6) $\forall y(y \in Y \Rightarrow \overline{R}ayc)$. Aber aus (3) und (6) folgt (7) $\overline{R}abc$ im Widerspruch zu (2).

RL: Wir nehmen an (1) die rechte Seite, sowie (2) $SaYc$, d.h. (Def. S) (3) $\exists y : y \in Y \;\&\; Rayc$. Dann folgt aus (1) und (3), daß $c \in Z$, wie gewünscht. ∎

[21] Im Kapitel über Konditionale haben wir $R_X ac$ statt $SaXc$ geschrieben. Die jetzige Notation soll einerseits den Unterschied zur dreistelligen Relation R wahren und andererseits die aufzuzeigende Beziehung zwischen $SaXc$ und $Rabc$ andeuten.

Wenn wir nun, umgekehrt, von Modellen $\mathcal{M}^S$ des Typs $(W, 0, S, I)$ ausgehen – wir wollen sie hier *Chellas-Modelle* nennen [22] –, dann geht das (beinahe) genauso. Jetzt definieren wir

(Def. R) $\qquad Rabc$ gdw $\forall Y : b \in Y \Rightarrow SaYc.$

Um den gewünschten Effekt in Chellas-Modellen zu beobachten, müssen wir diese unter zwei Bedingungen stellen.

(pref) $\qquad X \subseteq Y \;\&\; SaXb \Rightarrow SaYb,$

(disj) $\qquad Sa(X \cup Y)b \Rightarrow SaXb$ oder $SaYb.$

Äquivalent können wir die beiden Bedingungen auch zu einer einzigen verschmelzen:

$$Sa(X \cup Y)b \text{ gdw } SaXb \text{ oder } SaYb.$$

Die Bedingung (pref) entspricht in einem Sinne, den der Leser selbst nachrechnen möge, dem Präfixschema B. Die Bedingung (disj) sorgt in Modellen für Sprachen mit Disjunktion dafür, daß das Schema $(A \to C) \to ((B \to C) \to (A \vee B \to C))$ gültig ist; und umgekehrt erzwingt die Gültigkeit des Schemas in einem Modell, daß dieses (disj) erfüllt. In rein implikativen Sprachen gibt es kein Schema, dessen Gültigkeit die Modellbedingung (disj) erzwingt. Das kann uns hier aber nicht daran hindern, Chellas-Modelle mit (disj) zu betrachten, zumal rein implikative Sprachen für uns nur ein Trittstein zum vollen Vokabular sein sollen.

Die zuvor gemachte Beobachtung können wir jetzt an $\mathcal{M}^S$ wiederholen.

Beobachtung 19. *In $\mathcal{M}^S$ gilt für alle $a \in W$ und alle $Y, Z \subseteq W$: $(\forall z : SaYz \Rightarrow z \in Z)$ gdw $(\forall yz : Rayz \;\&\; y \in Y \Rightarrow z \in Z)$.*

Beweis. LR: Wir nehmen an (1) die linke Seite, sowie (2) $Rabc$ und (3) $b \in Y$. (2) bedeutet (Def. S): (4) $\forall Y : b \in Y \Rightarrow SaYc$. Aus (3) und (4) folgt (5) $SaYc$ und also, nach (1), $c \in Z$, wie gewünscht.

RL: Wir kontraponieren und nehmen an (1) $\exists z : SaYz \;\&\; z \notin Z$, ferner – für *reductio* – (2) $\neg\exists y : Rayz \;\&\; y \in Y$ ($z \notin Z$ ist schon in (1) angenommen), d.h. (3) $\forall y : y \in Y \Rightarrow \overline{R}ayz$. $\overline{R}ayz$ bedeutet (Def. R): $\exists X : y \in X \;\&\; \overline{S}aXz$. Also ist (3) äquivalent zu (4) $\forall y : y \in Y \Rightarrow (\exists X : y \in X \;\&\; \overline{S}aXz)$. Wir fassen alle in (4) postulierten Mengen X in der Menge M zusammen. Dann folgt aus (4), (5) $\forall X \in M : \overline{S}aXz$. Aus (5) folgt nach (disj) (6) $\overline{S}a(\bigcup M)b$. Nun ist aber $Y \subseteq \bigcup M$. Also aus (6) nach (pref), $\overline{S}aYz$ – im Widerspruch zu unserer Annahme (1). ∎

[22] Die Bezeichnung der Modelle soll an [52] erinnern. Dort wurde zum ersten Mal die Äquivalenz von $(S \to)$ und $(R \to)$ erwähnt. Der Zusammenhang mit Relevanzlogiken (und einer Reihe weiterer Logiken) ist in [90] dargestellt; vgl. auch [24].

Die beiden Beobachtungen geben den genauen Sinn an, in dem $(S \rightarrow)$ als auch $(R \rightarrow)$ zwei Arten sind, dieselbe allgemeine Form der Interpretation von Konditionalen zu repräsentieren. Für spezifische Konditionale, wie materiale, strikte oder variabel strikte, wird die Form unter Bedingungen gestellt, die Vereinfachungen ermöglicht. Diese Vereinfachungen können so einschneidend ausfallen, daß die Illusion einer grundsätzlichen Verschiedenartigkeit erzeugt wird. Tatsächlich fallen aber auch solche Konditionale in den Modellierungsbereich der ternären Semantik nach Routley und Meyer – nur eben in mehr oder weniger eklatant redundanter Weise.

An dieser Stelle wollen wir eine weitere Beobachtung anschließen. Chellas-Modelle unterscheiden sich von den Modellen für die Konditionallogik im Kapitel 4. Dort waren die Modelle vom Typ (W, S, I) mit der Interpretationsbedingung $(S \rightarrow)$; Wahrheit im Modell war Wahrheit an allen seinen Punkten. Hier sind die Modelle vom Typ $(W, 0, S, I)$, ebenfalls mit der Bedingung $(S \rightarrow)$; Wahrheit im Modell ist jetzt Wahrheit an seinem Nullpunkt. In den zwei soeben angestellten Beobachtungen spielte der Nullpunkt keine Rolle: Es ging um die Interpretation von Konditionalen an beliebigen Punkten. Aber für die Frage, welche Formeln gültig seien, macht die Beurteilung am Nullpunkt erwartungsgemäß einen großen Unterschied.

So ist die grundlegende Konditionallogik **CK** unter anderem durch die Regel

RN. $$\frac{B}{A \rightarrow B}$$

gekennzeichnet. In Modellen vom Typ (W, S, I) ((W, S)-Modelle) ist RN unausweichlich. Denn, wenn B an allen Punkten wahr ist, dann natürlich auch an allen Punkten, die eine beliebig gewählte Bedingung erfüllen. In $(W, 0, S)$-Modellen sieht das anders aus. Hier ist eine Regel gültig, wenn sie Wahrheit am Nullpunkt von den Prämissen auf die Konklusion überträgt. Unter der Annahme, daß $0 \models B$, bleibt aber offen, ob $0 \models A \rightarrow B$. Denn an einem Punkt a mit $S0[\![A]\!]a$ mag B durchaus falsch sein.

Für die Gültigkeit des Schemas

Id. $$A \rightarrow A$$

haben wir (W, S)-Modelle unter die Bedingung

$$SaXb \Rightarrow b \in X$$

gestellt. Diese Bedingung garantiert natürlich auch, daß Id in $(W, 0, S)$-Modellen gültig ist. Aber sie schießt über das Ziel hinaus, indem sie ebenso das relevanzlogisch nicht akzeptable Schema $A \to (B \to B)$ als gültig auszeichnet. Tatsächlich kann man sich schnell davon überzeugen, daß für Id die schwächere Bedingung

$$S0Xa \Rightarrow a \in X$$

hinreichend und notwendig ist. Diese Bedingung für $(W, 0, S)$-Modelle entspricht der Weitergabe-Bedingung in $(W, 0, R)$-Modellen. Für die Semantik der Relevanzlogik ist also nicht nur eine Bedingung wie $(R \to)$ oder $(S \to)$ wesentlich, sondern auch die Präsenz von nicht-normalen Punkten, d.h. solchen, an denen gültige Formeln falsch werden können.

6. Implikation und Negation

Zur im letzten Abschnitt beschriebenen Basistheorie der Implikation, $\mathbf{C}_{\to}$, wollen wir jetzt die Theorie der Negation aus dem Abschnitt über die DeMorgan-Logik hinzufügen. Dazu erweitern wir A-Rahmen um den Routley-Stern zu *Routley-Meyer-Rahmen*, hier *B-Rahmen* genannt,

$$(W, 0, {}^*, R).$$

Wie in Definition 3 erklärt, ist * eine Involution, d.h. es gilt

$$\text{(inv)} \qquad a = a^{**}.$$

Ferner werden wir die Bedingung

$$\text{(R4)} \qquad Rabc \Rightarrow Rac^*b^*$$

benötigen. Schließlich sollen B-Rahmen die Bedingungen der Reflexivität (R1), Monotonie (R2) und Assoziativität (R3) für die Relation R erfüllen, und in Modellen auf B-Rahmen fordern wir die Einhaltung der atomaren Weitergabe-Bedingung (Atm $\leq$). In *Modellen für* $\neg$ *und* $\to$ auf B-Rahmen werden Negation und Implikation so interpretiert:

$$(\neg) \qquad a \models \neg A \text{ gdw } a^* \not\models A;$$

$$(\to) \qquad a \models A \to B \text{ gdw } \forall xy : Raxy \;\&\; x \models A \Rightarrow y \models B.$$

Wie zuvor ist *Wahrheit in einem Modell* Wahrheit am Nullpunkt des Modells. Eine Formel ist genau dann *gültig* in der Klasse aller B-Rahmen, wenn sie in allen Modellen (für $\neg$ und $\to$) auf allen Rahmen dieser Klasse wahr ist. Die Menge der so gültigen Formeln bezeichnen wir als (Negations- und Implikations-) *Theorie der B-Rahmen*, $Th_{\neg\rightarrow}(\mathbb{B})$.

Anmerkung. Rahmen mit der Signatur $(W, 0, *, R)$ sind nicht die allgemeinsten Strukturen, die in der Modelltheorie von Relevanzlogiken verwendet werden. So werden in der Literatur auch sogenannte unreduzierte Rahmen des Typs $(W, O, 0, *, R)$ betrachtet, in denen 0 ein ausgezeichnetes Element einer Teilmenge O von W ist. Die Modellierungsbreite umfaßt so auch sehr schwache Logiken, die wir hier nicht behandeln; siehe insbesondere [254]. Eine Ausnahme stellt die Logik **DW** dar, für die wir am Ende dieses Kapitels ein Sequenzenkalkül beschreiben werden. Schemata wie WI (s.u.) erzwingen die Reduktion der Menge O zu einer Einermenge, woraufhin wir mit reduzierten Rahmen arbeiten können.

Wir werden in diesem Abschnitt zeigen, daß die folgenden Axiome und Regeln genau die Theorie der B-Rahmen axiomatisiert:

Das System $\mathbf{C}_{\neg\rightarrow}$

$\mathbf{C}_{\rightarrow}$ (p. 366) + ...

DNI. $A \rightarrow \neg\neg A$

DNE. $\neg\neg A \rightarrow A$

Kp′. $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$

Für die Axiomatisierung genügt nun das Präfix-Schema B, da das Suffix-Schema

CB. $(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$

sich aus B und Kp so herleiten läßt (etwas verkürzt):

(1) $(\neg B \rightarrow \neg A) \rightarrow ((\neg C \rightarrow \neg B) \rightarrow (\neg C \rightarrow \neg A))$ B
(2) $(\neg\neg A \rightarrow \neg\neg B) \rightarrow ((\neg\neg B \rightarrow \neg\neg C) \rightarrow (\neg\neg A \rightarrow \neg\neg C))$ (1), Kp′ etc.
(3) $(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$ (2), DNE etc.

Ensprechend wird auch die mit CB einhergehende Rahmenbedingung (R3*) in der Definition der B-Rahmen nicht mehr eigens benötigt.

Das Argument für die Behauptung, daß

$$\mathbf{C}_{\neg\to} = Th_{\neg\to}(\mathbb{B}),$$

ist im Prinzip so wie das Argument für die Logik $\mathbf{C}_{\to}$ im vorhergehenden Abschnitt. An einigen Stellen muß das Argument ergänzt werden. Für den Richtigkeitsteil der Behauptung werden zwei Ergänzungen nötig, deren Ausführungen dem Leser als einfache Übungen überlassen seien:

1. Für das *Weitergabelemma* 9 müssen wir im Induktionsschritt den Fall $A = \neg B$ betrachten. Dazu verwenden wir die neue Rahmenbedingung (R4).
2. Für den *Richtigkeitssatz* 12 müssen wir zeigen, daß die Schemata DNI, DNE und Kp′ gültig sind. Für die DN-Schemata verwenden wir (inv), für die Kontraposition Kp′ wird (R4) benötigt.

So kommen wir gleich zum Vollständigkeitsteil der Behauptung. Die Definition eines *kanonischen Modells* für $\mathbf{C}_{\neg\to}$erweitert die Definition für $\mathbf{C}_{\to}$ um eine Operation * so:

(Def *) $$A \in a^* \text{ gdw } \neg A \notin a.$$

Die Punkte in der kanonischen Modellmenge W sind, wie zuvor, 0-implikative Mengen von Formeln, wobei jetzt der Nullpunkt die Theoremmenge von $\mathbf{C}_{\neg\to}$ enthält.

Als erstes müssen wir uns davon überzeugen, daß die gerade definierte Operation * tatsächlich eine Abbildung von W auf W im kanonischen Modell für $\mathbf{C}_{\neg\to}$ ist, d.h. daß sie die Eigenschaft einer Formelmenge, 0-implikativ zu sein, bewahrt.

LEMMA 20. (Sternlemma) *Wenn a ein Punkt in einem kanonischen Modell (für $\mathbf{C}_{\neg\to}$) ist, dann ist auch a^* (nach Def. *) ein solcher Punkt.*

BEWEIS. Es genügt hier zu zeigen, daß a^* 0-implikativ ist unter der Annahme, daß a es ist. Dazu nehmen wir an, (1) $A \in a^*$ und (2) $A \to B \in 0$. Dann folgt aus (2) und Kp, (3) $\neg B \to \neg A \in 0$. Nun folgt aus (1) nach (Def *) (4) $\neg A \notin a$. Also aus (3), (4) und der Annahme, daß a eine Theorie ist, (5) $\neg B \notin a$, d.h. nach (Def *), daß $B \in a$, wie gewünscht. ∎

LEMMA 21. (Hauptlemma) *Jedes kanonische Modell für $\mathbf{C}_{\neg\to}$ erfüllt die Bedingungen für Modelle auf B-Rahmen.*

BEWEIS. *I.* Für die *Rahmenbedingungen* erweitern wir den Beweis von Lemma 13, indem wir (inv) und (R4) nachweisen. Für die Bedingung (inv) benutzen wir den Umstand, daß 0 alle Instanzen der DN-Schemata enthält; für (R4) greifen wir in gleicher Weise auf das Schema Kp zurück.

II. Für die *Modellbedingungen* bauen wir ebenfalls auf dem Beweis von Lemma 13 auf und stellen fest, daß ($\neg$) aufgrund der Definition (Def *) erfüllt ist. ■

Sei $F \notin \mathbf{C}_{\neg\to}$. Dann gibt es ein (kanonisches) Modell mit $0 \not\models F$. Zusammen mit der Richtigkeit von $\mathbf{C}_{\neg\to}$ in B-Modellen folgt so:

SATZ 22. (Richtigkeit und Vollständigkeit) $\mathbf{C}_{\neg\to} = Th_{\neg\to}(\mathbb{B})$.

Wir illustrieren den Gebrauch der Modelle indem wir zeigen, daß eine Formel *nicht* gültig ist. Dazu betrachten wir die folgende Instanz von EFQ in implikativer Form:

$$P \to (\neg P \to Q).$$

Die Formel wird genau dann in einem Modell falsch, wenn dieses Punkte $0, a, b, c, c^*, d$ enthält mit $R0ab$, $Rbcd$ und $a \models P$, $c \models \neg P$ sowie $d \not\models Q$. Klarerweise gibt es solche Modelle. Es sei zum Beispiel $a = b = c = d$. Wir betrachten ein Modell $(\{0, a, a^*\}; R0aa, Raaa, I)$ mit $I(P, a) = 1$, $I(P, a^*) = 0$ und $I(Q, a) = 0$. Dann:

W	*wahr*	*falsch*
0		$P \to (\neg P \to Q)$
a	$\neg P, P$	$Q, \neg P \to Q$
a^*	$\neg Q$	P

"Reine" Relevanzlogik. In der klassischen Logik ist die Theorie der Implikation auf die der Negation und Konjunktion bzw. der Negation und Disjunktion reduzierbar. Die materiale Implikation $A \supset B$ bedeutet danach nichts anderes als $\neg(A \wedge \neg B)$ bzw, $\neg A \vee B$. Das ist in der Relevanzlogik nicht der Fall, d.h. das richtige Verständnis der Implikation ergibt sich nicht definitorisch aus dem richtigen Verständnis der Junktoren $\neg$ und $\wedge$ bzw. $\neg$ und $\vee$ – jedenfalls nicht, wenn diese in klassischer Weise zu verstehen sind.

Wir können die Sache auch umgekehrt auffassen: Da die Menge $\{\neg, \supset\}$ funktional vollständig ist, so können wir auch die klassische Bedeutung der Disjunktion und Konjunktion aus den klassischen Bedeutungen von Negation und Implikation ableiten. Danach bedeutet $A \vee B$ dasselbe wie $\neg A \supset B$ und $A \wedge B$ ist synonym für $\neg(A \supset B)$.

Im Prinzip geht das auch in der Relevanzlogik: Bedeutungen von Disjunktion und Konjunktion können wir auch aus den relevanzlogischen Bedeutungen von Negation und Implikation ableiten. Nur sind die so definierten

Junktoren dann nicht die klassische, sondern die relevanzlogische Disjunktion und Konjunktion. Um nicht verschiedene Dinge gleich zu benennen, wollen wir die relevanzlogische Konjunktion bzw. Disjunktion – an die englischen Bezeichnungen angelehnt – mit *Fusion* bzw. *Fission* bezeichnen und dafür die Symbole $\circ$ bzw. $+$ verwenden.

Hier also sind die Definitionen von Fusion und Fission aus der relevanzlogischen Negation und Implikation:

Fusion : $\quad A \circ B := \neg(A \to \neg B),$

Fission : $\quad A + B := \neg A \to B.$

Die dazugehörigen Interpretationsklauseln sind

$(\circ) \qquad a \models A \circ B$ gdw $\exists xy : Rxya \;\&\; x \models A \;\&\; y \models B,$ und

$(+) \qquad a \models A + B$ gdw $\forall xy : Raxy \;\Rightarrow\; x^* \models A$ oder $y \models B.$

Die Fusion als eine Art Konjunktion zu bezeichnen, ist in gewisser Hinsicht irreführend. So steht die Fusion nicht unter der für die Konjunktion typischen Beseitigungsregel:

$$(*) \qquad A \circ B \to A$$

kann *kein* Theorem einer Relevanzlogik sein. Denn wir haben schon in $\mathbf{C}_{\neg\to}$ das Theorem

$$(A \circ B \to C) \to (A \to (B \to C)),$$

aus dem *per* MP mit $(*)$ das Nicht-Theorem $A \to (B \to A)$ folgen würde. Die naheliegendste Lesart von $(\circ)$ deutet die Fusion als einen Junktor, der die Verträglichkeit der Fusionsteile anzeigt: sie schließen einander nicht aus. Ähnliches gilt für die Interpretation der Fission als eine Art Disjunktion. Es gilt eben nicht $A \to A + B$, d.h. die für die Disjunktion typische Einführungsregel. Die Fission drückt vielmehr so etwas wie eine vollständige Partition von Möglichkeiten an: Wenn der eine Fissionsteil falsch ist, dann muß der andere wahr sein. [23]

Fusion und Fission können schwerlich ein eigenständiges Interesse beanspruchen. Was es darüber zu sagen gibt, sind Korollare zu Beobachtungen über die Implikation (und Negation). Jedoch sind diese Junktoren zuweilen willkommene Darstellungsmittel. In diesem Sinne werden wir am Ende dieses Kapitels vom Fusionsjunktor noch einmal Gebrauch machen.

---

[23] Vgl. die Diskussion dieser Junktoren in [10, pp. 176f u. 344ff] und insbesondere in [240, Kap. 3].

7. Das volle Vokabular – Die Logik C

In der Routley-Meyer-Semantik sind die Implikation und die Negation intensionale Junktoren in dem Sinne, daß die Wahrheit (oder Falschheit) einer Implikation bzw. Negation an einem Punkt auf die Wahrheit (oder Falschheit) ihrer Teilformeln an anderen Punkten verweist. Im typischen vollen Vokabular einer Relevanzlogik werden diese intensionalen Junktoren ergänzt um die extensionale Konjunktion und Disjunktion. Die Modelle werden nun so eingerichtet, daß diese vier Junktoren im Sinne der De-Morgan-Äquivalenzen zusammenwirken.

Zu diesem Zweck schränken wir die Rahmenklasse $\mathbb{B}$ durch eine neue Bedingung (R5) zur Rahmenklasse $\mathbb{C}$ ein und erweitern gleichzeitig die Definition eines Modells um die Bedingungen für die extensionalen Junktoren.

Definition 23. (C-Rahmen und -Modelle.)
Ein *C-Rahmen* ist ein Quadrupel $(W, 0, {}^*, R)$ so, daß W, 0, *, und R wie in einem B-Rahmen sind und die folgenden Bedingungen gelten:

(inv) $$a = a^{**}$$

(R1) $$a \leq a$$

(R2) $$a \leq b \ \&\ Rbcd \Rightarrow Racd$$

(R3) $$R^2(ab)cd \Rightarrow R^2a(bc)d$$

(R4) $$Rabc \Rightarrow Rac^*b^*$$

(R5) $$Raaa$$

Ein *Modell* $(W, 0, R, {}^*, I)$ auf einem C-Rahmen fügt diesem eine Bewertung $I : \mathrm{ATM} \longrightarrow \{0, 1\}$ hinzu unter der Bedingung:

(Atm $\leq$) $$I(P, a) = 1 \ \&\ a \leq b \Rightarrow I(P, b) = 1.$$

Ein Modell ist ein *Modell für* $\neg, \wedge, \vee$ und $\rightarrow$, wenn darin eine Wahrmacherrelation $\models \subseteq W \times \mathrm{FML}$ unter den folgenden Bedingungen steht:

(P) $$a \models P \text{ gdw } I(P, a) = 1$$

($\neg$) $$a \models \neg A \text{ gdw } a^* \not\models A$$

($\wedge$) $$a \models A \wedge B \text{ gdw } a \models A \text{ und } a \models B$$

($\vee$) $$a \models A \vee B \text{ gdw } a \models A \text{ oder } a \models B$$

($\rightarrow$) $$a \models A \rightarrow B \text{ gdw } \forall xy : Raxy \ \&\ x \models A \Rightarrow y \models B$$

Im folgenden wollen wir unter einem Modell immer ein Modell für die genannten Junktoren verstehen. Mit $\mathbb{C}$ bezeichnen wir die Klasse aller C-Rahmen. Die *Theorie von* $\mathbb{C}$, welche wir jetzt kurz mit $Th(\mathbb{C})$ statt mit

$Th_{\neg\to\wedge\vee}(\mathbb{C})$ notieren, ist die Menge der Formeln, die in allen Modellen auf Rahmen in $\mathbb{C}$ wahr sind (also die Menge der in $\mathbb{C}$ *gültigen* Formeln). (*Ende der Definition.*)

In der folgenden Tafel sind die zwei verschiedenen Rahmentypen (Signaturen), die Bedingungen unter denen diese stehen, die Junktoren, die in ihnen interpretiert werden und die dazu gehörigen logischen Theorien zusammengefaßt.

	A-Rahmen	**B-Rahmen**	**C-Rahmen**
Signatur	$(W, 0, R)$	$(W, 0, {}^*, R)$	$(W, 0, {}^*, R)$
Bedingungen	(R1-4)	(inv),(R1-4)	(inv),(R1-5)
Junktoren	$\to$	$\neg, \to$	$\neg, \to, \wedge, \vee$
Theorie	$Th_{\to}(\mathbb{A}) = \mathbf{C}_{\to}$	$Th_{\neg\to}(\mathbb{B}) = \mathbf{C}_{\neg\to}$	$Th(\mathbb{C}) = \mathbf{C}$

In den A-Rahmen ging es um eine reine Theorie der Implikation. Das Thema der B-Rahmen war die Ergänzung der Implikationstheorie um die reine Theorie der Negation so, wie wir sie aus den Routley-Rahmen kennen. Die C-Rahmen fügen dies nun alles zusammen.

Dieser Abfolge relationaler Rahmen hatten wir die Routley-Rahmen, mit Signatur $(W, {}^*)$, vorangeschickt. Auf solchen Rahmen hatten wir Modelle definiert, in denen die Junktoren $\neg$, $\wedge$ und $\vee$ rekursiv interpretiert wurden. (Das schließt übrigens nicht aus, daß wir in solchen Modellen auch Formeln der Gestalt $A \to B$ "interpretieren". Nur gibt es für solche Formeln keine rekursiven Bedingungen, d.h. sie werden im Prinzip wie strukturlose Atome behandelt.) Das Resultat war die DeMorgan-Logik **DML**, eine unter verschiedenen Perspektiven stabile Theorie des Zusammenspiels der Negation mit der Konjunktion und der Disjunktion. Wie auf p. 356 erklärt, können wir den Bestand gültiger **DML**-Folgerungen (mit endlichen Antezedenz-Mengen) auch als eine Menge erstufiger Implikationen darstellen. Im Kern geht es jetzt darum, diese erststufigen Implikationen der Theorie der B-Rahmen hinzuzufügen.

Bevor wir uns der Frage nach der Theorie der C-Rahmen zuwenden, wollen wir uns die neue Bedingung (R5), *Raaa*, genauer ansehen. Ein wenig fällt sie hier "aus dem Rahmen", denn die unmittelbare Aufgabe besteht ja nur darin B-Modelle so zu erweitern, daß sie Konjunktionen und Disjunktionen im DeMorgan'schen Sinne interpretierten können. Dazu sollte es keiner neuen Bedingung für die Relation R bedürfen.

Die Bedingung (R5) bewirkt, daß das Modus-Ponens-Schema

WI. $$A \wedge (A \to B) \to B$$

gültig ist. In kanonischen Modellen für **C** führt dies dazu, daß alle Punkte unter Implikation abgeschlossen sind, d.h. wenn $A \in a$ und $A \to B \in a$, dann $B \in a$. Für das Vollständigkeitsarguments, welches gleich vorgestellt werden soll, ist das eine wichtige Bedingung für kanonische Modelle. Diese Bedingung wird durch die Präsenz von WI in der Logik auf einfache Weise sichergestellt. Es gilt jedoch anzumerken, daß sich das Vollständigkeitsargument so verallgemeinern läßt, daß Logiken auch ohne WI in den Modellierungsbereich von C-Rahmen *ohne* die Bedingung (R5) fallen. Insbesondere gilt das für die Logik $\mathbf{B} = \mathbf{C}-$WI, die so als eigentliche Basislogik für die Routley-Meyer-Semantik angesehen werden kann.[24] Der Aufwand für eine solche Verallgemeinerung ist jedoch erheblich und würde den Rahmen dieses Kapitels sprengen. Also wollen wir WI hier einfach zuhilfe nehmen.

Wir beginnen mit der Beobachtung, daß die nötige Induktion für das Weitergabelemma 9 (und sein Korollar) sich für die Fälle von konjunktiven und disjunktiven Formeln problemlos fortsetzen läßt. Also halten wir fest, daß für beliebige Formeln A und B sowie Punkte a und b gilt:

Weitergabe: Wenn $a \models A$ und $a \leq b$, dann $b \models A$;

Verifikation: $0 \models A \to B$ gdw $[\![A]\!] \subseteq [\![B]\!]$.

Nun fügen wir der Theorie der B-Rahmen, d.h. dem System $\mathbf{C}_{\neg\to}$, Schemata hinzu, die wir mehr oder weniger direkt aus **DML** übernehmen können. Es wird sich gleich herausstellen, daß wir so eine Theorie der C-Rahmen erhalten.

Daß das Weitergabe-Lemma auch für **C** gilt und jedes Theorem von **C** in $\mathbb{C}$ gültig ist, d.h. in $Th(\mathbb{C})$ enthalten ist, läßt sich auf hinlänglich bekannte Weise einfach nachprüfen. Dabei machen wir für das Schema WI Gebrauch von der Rahmenbedingung (R5) Es gilt daher:

Satz 24. (Richtigkeit) *Jedes Theorem von* **C** *ist in der Klasse* $\mathbb{A}$ *aller Rahmen gültig:* $\mathbf{C} \subseteq Th(\mathbb{A})$.

Für den Beweis der Umkehrung des Satzes, d.h. der Vollständigkeit, müssen wir etwas mehr Aufwand treiben als bei den Vollständigkeitssätzen für Fragmente von **C** ohne die extensionalen Junktoren $\wedge$ und $\vee$. Das mag zunächst erstaunen, da doch $\wedge$ und $\vee$ alte Bekannte sind, deren Bedeutung sich, im Gegensatz zu $\neg$ und $\to$, in einfachen Wahrheitstafeln angeben läßt.

---

[24] Die Basislogik **B** wird insbesondere prominent in [254] präsentiert. Dort wird sie als Theorie unreduzierter Modelle dargestellt; vgl. Fußnote 16, p. 365. Schon den Autoren von [254] war bekannt, daß für die Modellierung von **B** im Prinzip auch reduzierte Modelle ausreichen. In [268] wird der dafür nötige Aufwand im Detail beschrieben.

Das System **C**

$$\mathbf{C}_{\neg\rightarrow} \text{ (p. 380) } + \ \ldots$$

WI.	$A \wedge (A \rightarrow B) \rightarrow B$
Konj.	$A \wedge B \rightarrow A \qquad A \wedge B \rightarrow B$
Komp∧.	$((A \rightarrow B) \wedge (A \rightarrow C)) \rightarrow (A \rightarrow B \wedge C)$
Disj.	$A \rightarrow A \vee B \qquad B \rightarrow A \vee B$
Komp∨.	$(A \rightarrow C) \wedge (B \rightarrow C) \rightarrow (A \vee B \rightarrow C)$
Dist.	$A \wedge (B \vee C) \rightarrow (A \wedge B) \vee (A \wedge C)$
DM1.	$\neg(A \wedge B) \rightarrow \neg A \vee \neg B$
DM2.	$\neg A \wedge \neg B \rightarrow \neg(A \vee B)$
ADJ.	$\dfrac{A \quad B}{A \wedge B}$

Das Problem besteht darin, daß wir von den Punkten in kanonischen Modellen jetzt mehr fordern müssen, als daß sie unter Implikation am Nullpunkt abgeschlossen sind. Diese müssen nun auch zu ($\wedge$) und ($\vee$) analoge Bedingungen erfüllen, nämlich:

(konj) $\quad A \wedge B \in a$ gdw $A \in a$ und $B \in a$, und

(disj) $\quad A \vee B \in a$ gdw $A \in a$ oder $B \in a$.

Insbesondere die letzte Bedingung, die Disjunktionseigenschaft für Theorien, ist problematisch. Denn unter den Axiomen für **C** gibt es keine unmittelbar ersichtlichen Prinzipien, die garantieren, daß Formelmengen, die unter in **C** beweisbarer Implikation abgeschlossen sind, die Disjunktionseigenschaft besitzen. Formelmengen mit der Links-nach-Rechts-Richtung der Disjunktionseigenschaft nennen wir kurz *prim.*[25] Prime Formelmengen, die gewisse weitere Bedingungen erfüllen sollen, müssen mit Bedacht konstruiert werden. Das wird uns im Laufe des Vollständigkeitsarguments

[25] Die umgekehrte Richtung werden Theorien, die unter Implikation in **C** abgeschlossen sind aufgrund der Disj-Schemata erfüllen.

beschäftigen. Zunächst wieder einige Definitionen, von denen uns die erste schon bekannt ist. Es sei S eine Formelmenge.

- Eine Formelmenge X ist *S-implikativ*, wenn gilt:
  Wenn $A \in X$ und $A \to B \in S$, dann $B \in X$.
- X ist *adjunkiv*, wenn gilt: Wenn $A \in X$ und $B \in X$, dann $A \wedge B \in X$.
- X ist *prim*, wenn gilt: $A \vee B \in X$ gdw $A \in X$ oder $B \in X$.

Die ersten beiden Definitionen wollen wir auch so zusammenfassen: X ist eine *S-Theorie*, wenn X adjunktiv und S-implikativ ist.

Für ein beliebig gewähltes Nicht-Theorem F wollen wir ein **C**-kanonisches Modell $(W, 0, R, I)$ konstruieren, in dem W eine Menge primer **C**-Theorien und 0 ($\in W$) zudem eine Erweiterung von **C** ist. In der Nullpunktmenge soll F nicht enthalten und so falsifiziert sein. Wir beginnen mit der Konstruktion von 0.

Lemma 25. (Lindenbaum-Lemma für prime Erweiterungen) *Es sei X ($\supseteq$ **C**) eine **C**-Theorie und F sei eine Formel so, daß $F \notin X$. Dann gibt es eine Erweiterung $X'(\supseteq X)$ so, daß $F \notin X'$ und X' eine prime **C**-Theorie ist.*

Beweis. (Vergleiche den Beweis im Einleitungskapitel.) Wir gehen von einer Aufzählung $A_1, A_2, ...$ aller Formeln aus und konstruieren X' wie folgt. Dabei sei $\vdash$ die übliche Relation der Ableitbarkeit aus Prämissen, hier im System **C**. Es sei an folgende Eigenschaften (Reflexivität, Monotonie, Schnitt, Disjunktion) erinnert:

$$A \vdash A \qquad \frac{X \vdash A}{X, Y \vdash A} \qquad \frac{X, A \vdash B \quad X \vdash B}{X \vdash B} \qquad \frac{X, A \vdash C \quad X, B \vdash C}{X, A \vee B \vdash C}$$

Die ersten drei gelten für jede Ableitbarkeitsrelation, die letzte im wesentlichen aufgrund des Axioms Komp$\vee$.

$$\begin{aligned} X_0 &= X \\ X_{n+1} &= \begin{cases} X_n \cup \{A_{n+1}\}, & \text{falls } X_n, A_{n+1} \nvdash F; \\ X_n & \text{anderenfalls} \end{cases} \\ X' &= X_0 \cup X_1 \cup X_2 \cup \cdots \end{aligned}$$

(i) $X \subseteq X'$ nach Konstruktion.

(ii) $F \notin X'$. Wäre das nicht so, dann müßte F entweder schon in X sein, was nach Voraussetzung ausgeschlossen ist, oder es müßte ein X_i in der Konstruktion geben so, daß $X_i, F \nvdash F$, im Widerspruch zu Reflexivität und Monotonie.

(iii) X' *ist eine prime* **C**-*Theorie.* Denn X' ist ...

... **C**-*implikativ*. Denn, angenommen (1) $A \in X'$ und (2) $\vdash A \to B$. Wir nehmen ferner an – für *reductio* –, (3) $B \notin X'$. Dann gibt es in der Konstruktion einen Schritt i mit $X_i, B \vdash F$ und also, nach Monotonie, (4) $X', B \vdash F$. Aus (1) folgt (5) $X' \vdash A$. Ferner haben wir (6) $A, A \to B \vdash B$. Also (Schnitt) aus (5) und (6), (7) $X', A \to B \vdash B$. Aber $A \to B \in X'$ ($\supseteq X \supseteq \mathbf{C}$). Also (8) $X' \vdash B$. Nun folgt aus (4) und (8) (Schnitt), daß $X' \vdash F$ – im Widerspruch zu (ii).

... *adjunktiv.* Denn, angenommen (1) $A, B \in X'$, jedoch (2) $A \wedge B \notin X'$. Dann (3) $X', A \wedge B \vdash F$. Da aber nach (1) $X' \vdash A$ und $X' \vdash B$, so folgt (4) $X' \vdash A \wedge B$. Also (Schnitt) aus (3) und (4): $X' \vdash F$ – im Widerspruch zu (ii).

... *prim.* Denn, angenommen (1) $A \vee B \in X'$, und also $X' \vdash A \vee B$, aber – *reductio* – (2) weder $A \in X'$ noch $B \in X'$. Wie zuvor folgt aus (2), (3) $X', A \vdash F$ und $X', B \vdash F$. Also nach der Disjunktionsregel für $\vdash$, (4) $X', A \vee B \vdash F$. Aus (1) und (4) (Schnitt) folgt dann $X' \vdash F$ – im Widerspruch zu (ii). ∎

Wenn wir nun ein beliebiges Nicht-Theorem F von **C** betrachten, dann erfüllt die Theorie **C** natürlich selbst die Bedingung des Lemmas. (Denn aufgrund der Regeln MP und ADJ ist **C** selbstimplikativ und adjunktiv.) Also läßt sich für jedes Nicht-Theorem eine prime **C**-Theorie 0 konstruieren so, daß $\mathbf{C} \subseteq 0$ und $F \notin 0$. Damit haben wir den Nullpunkt eines kanonischen Modells, das F widerlegt. Die Menge W des kanonischen Modells soll nun die Menge aller primen 0-Theorien sein. Im Gegensatz zu 0 muß eine 0-Theorie nicht die Menge der **C**-Theoreme enthalten.

Aus der Existenz von 0 folgt leider nicht unmittelbar die Existenz primer 0-Theorien. Das Argument für die Existenz solcher Mengen ist jedoch im Prinzip genauso wie das für das vorangegangene Lemma, weshalb wir hier nur auf den einzigen wesentlichen Unterschied aufmerksam machen. Die Konstruktion beginnt mit einer beliebigen 0-Theorie. Der Schritt ist nun so zu definieren:

$$X_{n+1} = \begin{cases} X_n \cup \{A_n\}, \text{ falls } \neg\exists B_0, ..., B_m \in X_n \cup \{A_{n+1}\} \text{ so,} \\ \qquad \text{daß } B_0 \wedge \cdots \wedge B_m \to F \in 0; \\ X_n \text{ anderenfalls} \end{cases}$$

Wir können uns dann davon überzeugen, daß X' 0-implikativ, adjunktiv und prim ist. Dazu benutzen wir im wesentlichen die **C**-Theoreme

$(A \to C) \to (A \wedge B \to C)$,
$A \to (B \to A \wedge B)$, sowie

$(A \to C) \wedge (B \to C) \to (A \vee B \to C)$ (Komp$\vee$).[26]

Ein *kanonisches Modells für* $\mathbf{C}$,

$$(W^{\mathbf{C}}, 0^{\mathbf{C}}, *^{\mathbf{C}}, R^{\mathbf{C}}, I^{\mathbf{C}}),$$

besteht also aus

1. W, einer nichtleeren Menge primer 0-Theorien;
2. $0 \in W$ und $\mathbf{C} \subseteq 0$. Die übrigen Definitionen sind wie zuvor:
3. $^* : W \longrightarrow W$ so, daß $A \in a^*$ gdw $\neg A \notin a$;
4. $R \subseteq W^3$ so, daß
   $Rabc$ gdw für alle Formeln A und B: $A \to B \in a$ & $A \in b \Rightarrow B \in c$;
5. $I : \mathrm{ATM} \times W \longrightarrow \{0, 1\}$ so, daß $I(P, a) = 1$ gdw $P \in a$.
   Schließlich sei $a \models A$ gdw $A \in a$.

Das nächsten Lemma stellt sicher, daß die Zeile 3 in der Definition eines kanonischen Modells nicht zu einer leeren Definition führt.

Lemma 26. (Sternlemma für adjunktive und prime Punkte) *Wenn a ein Punkt in einem kanonischen Modell (für* $\mathbf{C}$*) ist, dann ist auch a^* ein solcher Punkt.*

Beweis. Im Beweis von Lemma 20 wurde bereits gezeigt, daß a^* 0-implikativ ist, wenn a es ist. Wir ergänzen den Beweis jetzt wie folgt.

1. Wenn $a \in W$, dann ist a^* adjunktiv. Denn, angenommen (1) $A \in a^*$ und $B \in a^*$. Dann ist aufgrund der Def. von $*$, (2) $\neg A \notin a$ und $\neg B \notin a$. Da a prim ist, folgt aus (2), (3) $\neg A \vee \neg B \notin a$. Aber (DM1!) (4) $\neg(A \wedge B) \to \neg A \vee \neg B \in 0$. Also folgt aus (3) und (4), (5) $\neg(A \wedge B) \notin a$, d.h. $A \wedge B \in a^*$.

2. Wenn $a \in W$, dann ist a^* prim. Denn, angenommen (1) $A \vee B \in a^*$, d.h. (2) $\neg(A \vee B) \notin a$. Da nach DM2 (3) $\neg A \wedge \neg B \to \neg(A \vee B) \in 0$, so folgt aus (2), (4) $\neg A \wedge \neg B \notin a$. Da a adjunktiv ist, so folgt aus (4), (5) $\neg A \notin a$ oder $\neg B \notin a$, d.h. (nach Def. $*$) $A \in a^*$ oder $B \in a^*$. ∎

Lemma 27. (Hauptlemma) *Jedes kanonische Modell für* $\mathbf{C}$ *erfüllt die Bedingungen für Modelle auf C-Rahmen.*

Beweis. Wir bauen wieder auf dem Beweis des Hauptlemmas 13 für $\mathbf{C}_{\to}$ (und danach für $\mathbf{C}_{\neg\to}$) auf. Punkte im $\mathbf{C}$-kanonischen Modell müssen stärkere Bedingungen erfüllen als die in den zuvor betrachteten kanonischen Modellen. Deshalb gibt es zwei Sorten von Ergänzungen: Erstens müssen neue Bedingungen verifiziert werden; zweitens, müssen wir noch einmal jene

[26] Die Skizze beschreibt den hier einschlägigen Fall des sogenannten paarweisen Erweiterungslemmas in [11, §42.2].

Teile des Beweises prüfen, in denen die Existenz von Punkten im Modell behauptet wurde – erfüllen die konstruierten Punkte auch die stärkeren Bedingungen?

I. Rahmenbedingungen.

Ad (R5): Angenommen, (1) $A \to B \in a$ und (2) $A \in a$. (Zu zeigen: $B \in a$.) Da $a \in W$ adjunktiv ist, folgt aus (1) und (2), (3) $A \wedge (A \to B) \in a$. Der Nullpunkt enthält alle Instanzen des Schemas WI, d.h. (4) $A \wedge (A \to B) \to B \in a$. Da der Punkt a unter Implikation in 0 abgeschlossen ist, folgt aus (3) und (4), daß $B \in a$.

Noch einmal *ad* (R3): Im Konsequens von (R3) wird die Existenz eines Punktes mit bestimmten Eigenschaften gefordert. Im Beweis von Lemma 13 wurde ein Punkt definiert, der die in kanonischen Modellen für $\mathbf{C}_{\to}$ geforderte Eigenschaft der 0-Implikativität erfüllt. Nun müssen wir zusätzlich zeigen, daß der definierte Punkt auch adjunktiv und prim ist.

Man schlage die Definition (Dy) auf p. 368 nach. Dort wurde unter Bezug auf Punkte a,b,c, und d ein neuer Punkt y definiert. Von den Punkten a,b,c und d wollen wir jetzt annehmen, daß sie in W sind, d.h.

1. a,b,c und d sind 0-implikativ, adjunktiv und prim.

Wie zuvor können wir unter der Annahme $R^2(ab)cd$ feststellen:

2. y ist 0-implikativ;
3. $Rbcy$;
4. $Rayd$.

Ferner gilt auch

5. y ist adjunktiv.

Aus (Dy) läßt sich jedoch nicht ableiten, daß y, wie erforderlich, prim ist. Hier hilft das gleich zu beweisende Primlemma 28:

(P) Wenn $Rayd$ mit a, y und d wie in 1, 2 und 5, dann gibt es einen (primen) Punkt $y' \in W$ so, daß $y \subseteq y'$ und $Ray'd$.

Mithilfe von (P) folgt so, daß es einen Punkt $y' \in W$ gibt so, daß $Ray'd$. Aus 3 und $y \subseteq y'$ folgt ferner, daß $Rbcy'$. Also haben wir $R^2a(bc)d$, wie für (R3) gefordert.

II. Modellbedingungen.

Ad ($\neg$): Nach dem Sternlemma 26 wie für $\mathbf{C}_{\neg\to}$-kanonische Modelle.

Ad ($\wedge$) *et* ($\vee$): Da Punkte im kanonischen Modell jetzt adjunktiv und prim sind, ist dies für jeweils eine Richtung trivial. Für die anderen Richtungen benutzen wir Konj und Disj.

Noch einmal *ad* ($\to$): In der Rechts-nach-Links-Richtung ist zu zeigen:

$$(*) \qquad A \to D \notin a \;\Rightarrow\; \exists xy : Raxy \;\&\; A \in x \;\&\; D \notin y.$$

Dazu haben wir zwei Punkte b und c definiert – siehe (Db-c) auf p. 369 –, die zwar die im Konsequens von ($*$) geforderten Eigenschaften haben (b für x, c für y), jedoch nicht prim sein müssen. Für c ist das Problem leicht zu beheben. Wir verfahren so, wie im Anschluß an Lemma 25 geschildert, d.h. wir erweitern die 0-Theorie c zu eine einer primen 0-Theorie $c' \in W$. Dann ist folgendes der Fall:

1. a und c' sind in W;
2. b ist 0-implikativ und adjunktiv;
3. $Rabc'$;
4. $A \in b$;
5. $D \notin c'$.

Der Punkt c' kann also für y in ($*$) eintreten. Auch b könnten wir so wie c zu einem Punkt $b' \in W$ erweitern, hätten dann aber keine Garantie, daß $Rab'c'$. Auch hier hilft das Primlemma 28:

(P) Wenn $Rabc'$ und a, b und c' wie in 1-2, dann gibt es einen Punkt $b' \in W$ so, daß $b \subseteq b'$ und $Rab'c'$.

Aus 1-5 und (P) folgt nun unmittelbar, daß das Konsequens von ($*$) erfüllt ist. Damit endet der Beweis des Hauptlemmas. ∎

Es ist an der Zeit, den Beweis von (P) nachzureichen. Das folgende Lemma (auch "*Squeeze*-Lemma" genannt) ist tatsächlich etwas stärker als (P), denn die Annahme, daß die erste Koordinate von R prim ist, wird nicht benötigt.

Lemma 28. (Primlemma) *Wenn $Rabc$ mit a, b und c 0-implikativ und adjunktiv und c darüberhinaus prim ist, dann gibt es eine Menge b' so, daß $b \subseteq b'$, $Rab'c$ und b' 0-implikativ, adjunktiv und prim ist.*

Beweis. Wir definieren

$$Y := \{C : \exists B : C \to B \in a \;\&\; B \notin c\}$$

und beobachten:

1. Y ist unter Disjunktion abgeschlossen. Denn, angenommen (1) $C_1, C_2 \in Y$. Dann folgt nach der Definition, daß es C_1 und C_2 gibt mit (2) $C_1 \to B_1, C_2 \to B_2 \in a$ und (3) $B_1, B_2 \notin c$. Da a adjunktiv und 0-implikativ ist, folgt aufgrund des **C**-Theorems

$$(C_1 \to B_1) \wedge (C_2 \to B_2) \to (C_1 \vee C_2 \to B_1 \vee B_2),$$

daß (4) $C_1 \vee C_2 \rightarrow B_1 \vee B_2 \in a$. Da c prim ist, folgt aus (3), daß (5) $B_1 \vee B_2 \notin c$. Aber dann folgt aus (4) und (5) nach der Definition, (6) $C_1 \vee C_2 \in Y$.

2. $b \cap Y = \emptyset$. Denn, angenommen – für *reductio* – (1) $C \in b$ und (2) $C \in Y$. Letzteres bedeutet nach der Definition, daß es ein B gibt mit (3) $C \rightarrow B \in a$ und (4) $B \notin c$. Wir arbeiten unter der Annahme, daß (5) $Rabc$. Aber dann folgt aus (3), (1) und (5), daß $B \in c$, im Widerspruch zu (4).

3. X sei adjunktiv und 0-implikativ, Y sei unter Disjunktion abgeschlossen, und $X \cap Y = \emptyset$. Dann gibt es eine adjunktive, 0-implikative und prime Menge $X' \supseteq b$ mit $X' \cap Y = \emptyset$. Das ist eine Verallgemeinerung der Beobachtung, die wir im Anschluß an Lemma 25 gemacht haben. Im dort skizzierten Beweis übernimmt Y die Rolle von F. Im Schritt muß dazu die Bedingung für $X_n + 1 = X_n \cup \{A_n\}$ so lauten:

$$\neg\exists B_0, ..., B_m \in X_n \cup \{A_n\} \text{ so, daß } \exists C_0, ..., C_m \in Y \\ \text{mit } B_0 \wedge \cdots \wedge B_m \rightarrow C_0 \vee \cdots \vee C_k \in 0.$$

In die Beobachtung 3 können wir b (für X) und Y so, wie in 1 und 2 beschrieben, einsetzen. Dann folgt:

4. Es gibt eine prime 0-Theorie b' so, daß $b \subseteq b'$ und $b' \cap Y = \emptyset$.

Es bleibt zu zeigen, daß

5. $Rab'c$. Angenommen, (1) $A \rightarrow B \in a$ und (2) $A \in b'$. Dann, aufgrund von 2, (3) $A \notin Y$. Aber daraus folgt nach der Definition, daß (4) $A \rightarrow B \in a \Rightarrow B \in c$. Also, aus (1) und (4), $B \in c$, wie gewünscht. ■

Korollar 29. (P) *Es gilt in kanonischen Modellen für* **C**: *Wenn* $Rabc$ *mit* $a, c \in W$ *und* b *eine 0-Theorie ist, dann gibt es einen Punkt* $b' \in W$ *so, daß* $b \subseteq b'$ *und* $Rab'c$.

Damit ist das Argument für die Vollständigkeit von **C** bezüglich der Klasse aller C-Rahmen beendet und wir dürfen festhalten:

Satz 30. (Richtigkeit und Vollständigkeit) $\mathbf{C} = Th(\mathbb{C})$.

8. Das Standardsystem R

Ähnlich wie im Falle der Modallogik, so sind auch unter dem Titel "Relevanzlogik" eine Vielzahl logischer Systeme im Umlauf. Im Falle der Modallogik findet das eine naheliegende Erklärung in dem Umstand, daß der zentrale Operator – die Kiste $\Box$ – sehr verschiedene Interpretationen und daher sehr verschiedene Eigenschaften haben kann. Modallogiken sind Versuche mehr oder weniger interessante Satzoperationen in einer Erweiterung der klassischen Logik zu erfassen: verschiedene Operationen – verschiedene Erweiterungen.

Im Falle der Relevanzlogik ist eine solche Erklärung bedeutend weniger naheliegend. In einem trivialen Sinne behandeln verschiedene Relevanzlogiken natürlich verschiedene Operationen der Implikation. (Quine: "Wer die Logik wechselt, wechselt das Thema.") Aber Relevanzlogik wird meist als Alternativprogramm zur klassischen Logik vorgestellt. Die klassische Logik, so wird gesagt, biete eine falsche Theorie der Implikation an. Das lädt unweigerlich die Frage ein, welche Relevanzlogik *die* richtige Theorie anbietet. Auf diese Frage gibt es keine systematische Antwort. Vielleicht ist sie auch falsch gestellt. In den abschließenden Betrachtungen zum Kapitel werden wir auf sie zurückkommen.

Unbeschadet der gerade angestellten Überlegungen, gibt es doch *de facto* so etwas wie ein Hauptsystem der Relevanzlogik. Dies ist das System **R**. Anderson und Belnap (in [10, pp. 349–352]) nennen nicht weniger als acht Gründe, warum **R** sich dieser Auszeichnung erfreuen darf. Nicht einmal gestandene Relevanzlogiker wie Routley scheinen diese Gründe besonders zu beeindrucken – und einige von ihnen sind wohl auch nicht ganz ernst gemeint.[27] Die vermutlich stärksten Gründe sind die zwei von Anderson und Belnap zuerst genannten. Erstens hat **R** – als das historisch älteste System der relevanten Implikation – gewissermaßen den Test der Zeit bestanden. Zweitens schält sich zumindest das implikative Fragment von **R** auf recht natürliche Weise aus Betrachtungen über den Gebrauch von Annahmen, d.h. deren Relevanz für die Ableitung einer Konklusion in einem System des Natürlichen Schließens, heraus. (Darüber gleich mehr.) Man könnte, drittens, noch hinzufügen, daß **R** eine sehr starke Relevanzlogik ist, die der klassischen Logik recht nahe kommt. Es ist daher naheliegend sie zu verwenden, um aus Annahmen bzw. themenspezifischen Axiomen starke

[27] Zu den nicht ganz ernst gemeinten Gründen gehört z.B. "Reichtum" (*richness*), womit euphemistisch die Unentscheidbarkeit von **R** gemeint ist. — Routley schreibt: "Without much in the way of justification, **R** has become known as 'the system of relevant implication'." [254, p. 290] Routley befürwortet in seinem Werk eine "pluralistische" Antwort auf die Frage nach "der" richtigen Logik und versucht zu zeigen, daß für viele Zwecke auch eine sehr schwache Logik eine gute Wahl sein kann.

Das System **R**

Id.	$A \to A$
B.	$(A \to B) \to ((C \to A) \to (C \to B))$
C.	$(A \to (B \to C)) \to (B \to (A \to C))$
W.	$(A \to (A \to B)) \to (A \to B)$
Konj.	$A \wedge B \to A \qquad A \wedge B \to B$
Komp$\wedge$.	$((A \to B) \wedge (A \to C)) \to (A \to B \wedge C)$
Disj.	$A \to A \vee B \qquad B \to A \vee B$
Komp$\vee$.	$(A \to C) \wedge (B \to C) \to (A \vee B \to C)$
Dist.	$A \wedge (B \vee C) \to (A \wedge B) \vee (A \wedge C)$
DNE.	$\neg\neg A \to A$
Kp.	$(A \to \neg B) \to (B \to \neg A)$
ADJ.	$\dfrac{A \quad B}{A \wedge B}$
MP.	$\dfrac{A \quad A \to B}{B}$

Theorien zu generieren und diese mit den entsprechend klassisch generierten Theorien zu vergleichen. Als Beispiel sei hier der Abschluß unter **R** der Dedekind-Peano-Axiome für die natürlichen Zahlen genannt, die sogenannte Relevante Arithmetik [80].

Die hier angegebene Axiomatisierung von **R** erweitert in einfacher Weise die für das System **C**. Wir wollen nun beispielhaft einige in diesem Zusammenhang wichtige Formelschemata und deren Beziehungen zueinander betrachten. Die Ableitungen sind teils skizzenhaft, Anwendungen von Modus Ponens werden nicht explizit erwähnt, “Trans.” deutet eine Kombination von CB und MP an. In jedem Fall sind die überschlagenen Schritte leicht nachzutragen.

1. In **R** (wie schon im implikativen Teilsystem **BCI**) sind das Permutations-Schema C und das Ablösungsschema CI äquivalent. (Siehe p. 375.)

2. In **R** ist das Importations-Schema ableitbar:

$$\text{W}'. \qquad (A \to (B \to C)) \to (A \wedge B \to C).$$

Wir stellen zunächst fest, daß aufgrund von I und B (schon in **C**) ein Schema konjunktiver Monotonie ableitbar ist:

$$\text{Mon}\wedge. \qquad (A \to C) \to (A \wedge B \to C).$$

Dann geht eine Ableitung von W′ so:

(1)	$(A \to (B \to C)) \to (A \to (A \wedge B \to C))$	Mon∧, B
(2)	$(A \to (A \wedge B \to C)) \to (A \wedge B \to (A \wedge B \to C))$	1, Mon∧
(3)	$(A \wedge B \to (A \wedge B \to C)) \to ((A \wedge B \to C)$	W
(4)	$(A \to (B \to C)) \to (A \wedge B \to C)$	1,2,3 Trans.

Schon in **C** sind W und W′ äquivalent und können daher in der Axiomatisierung von **R** füreinander eintreten, wie die folgende Ableitung zeigt:

(1)	$(A \to (A \to B)) \to (A \wedge A \to B)$	W′
(2)	$A \leftrightarrow A \wedge A$	Konj; Komp∧, Id, ADJ
(3)	$(A \to (A \to B)) \to (A \to B)$	1,2, Subst.

3. In **R** ist WI ableitbar.

(1)	$(A \to ((A \to B) \to B)) \ \to \ (A \wedge (A \to B) \to B)$	W′
(2)	$A \wedge (A \to B) \to B$	1, C

4. Das Suffix-Schema CB folgt aus dem Präfix-Schema B und C allein, d.h. ohne Umweg über Kontraposition, Kp. Der findige Beweis von Sobociński ist in [10, p. 80] wiedergegeben.

5. In **R** ist das Reductio-Schema ableitbar:

$$\text{Red.} \qquad (A \to \neg A) \to \neg A.$$

(1)	$A \to ((A \to \neg A) \to \neg A)$	C
(2)	$((A \to \neg A) \to \neg A) \ \to \ (A \to \neg(A \to \neg A))$	Kp
(3)	$A \to (A \to \neg(A \to \neg A))$	1,2 Trans.
(4)	$A \to \neg(A \to \neg A)$	3, W
(5)	$(A \to \neg A) \to \neg A$	Kp

Im Beweis haben wir von C und W Gebrauch gemacht. Es läßt sich zeigen, daß der Rückgriff auf diese Schemata notwendig ist. Das bedeutet, daß in Logiken, in denen C und W keine Theoreme sind – wie zum Beispiel **C** –, Reductio ein unabhängiges Prinzip ist. Wenn wir Reductio in einer Logik haben wollen, die schwächer ist als **R**, dann müssen wir das Prinzip explizit hinzufügen (bzw. in anderer Weise dafür sorgen, daß es ableitbar ist).

6. Die Einführung doppelter Negationen, DNI $A \to \neg\neg A$, folgt aus Id und Kp per MP. Ähnlich können wir die alternative Version der Kontraposition, Kp′ $(A \to B) \to (\neg B \to \neg A)$, ableiten. Im Kontext der übrigen Axiome sind $\{\mathrm{DNI}, \mathrm{Kp}\}$ und $\{\mathrm{DNI}, \mathrm{DNE}, \mathrm{Kp}'\}$ austauschbar.

* * *

Aus der Perspektive der Routley-Meyer-Semantik ist **R** keine natürlich ausgezeichnete Logik. Sie ist einfach die Theorie solcher C-Rahmen, in denen es auf die Reihenfolge der ersten zwei Koordinaten der R-Relation nicht ankommt:

$$Rabc \;\Rightarrow\; Rbac. \tag{c}$$

Dieser Bedingung entspricht das Axiom

C. $$A \to ((A \to B) \to B).$$

Satz 31. (Richtigkeit und Vollständigkeit) $\mathbf{R} = Th(c(\mathbb{C}))$.

Beweis. Um das entprechende Resultat für **C** *vis-à-vis* $\mathbb{C}$ zu ergänzen, genügt es die Korrespondenz zwischen C und (c) nachzuweisen.

1. Jedes C-Modell unter der Bedingung (c) macht das Schema C wahr – folgt unmittelbar aus (c).
2. Jedes kanonische Modell für **R** erfüllt die Bedingung (c). Denn, angenommen (1) $Rabc$, (2) $A \to B \in b$ und (3) $A \in a$. (Zu zeigen: $B \in c$.) Nun haben wir C in 0, und da a 0-implikativ ist, folgt aus (3), (4) $(A \to B) \to B \in a$. Aus (1), (4) und (2) folgt schließlich $B \in c$. ∎

9. Deduktionstheoreme in der Relevanzlogik

Obgleich auch die Negation in der Relevanzlogik eine besondere, nichtklassische Interpretation erfährt, so ist doch der hauptsächlich interessierende Junktor der Relevanzlogik die Implikation. Die ganze Unternehmung ist, zumindest traditionell, der Versuch, eine logische Theorie des Pfeils $\to$ als objektsprachliche Wiedergabe eines "echt" folgernden Wenn-Danns zu entwickeln.

Nun gehört zu jedem axiomatischen System – also auch zu **R** – eine Relation $\vdash$ der Ableitbarkeit aus Prämissen, die üblicherweise so eingeführt wird (vgl. Kap. I, §2.3): Eine *Ableitung* (in einem gegebenen System) von A aus einer Prämissenmenge X ist eine Folge

$$\langle A_0, ..., A_n \rangle$$

mit $A = A_n$, in der jede der Formeln A_i $(0 \leq i \leq n)$ eine von drei Bedingungen erfüllen muß:

1. A_i ist ein Axiom.
2. A_i ist in X.
3. A_i ist Konklusion einer der Regeln (des Systems), wobei die Prämissen der Regelanwendung dem A_i in der Folge vorangehen.

Wenn wir auf diese Weise $\vdash$ auf der Basis von **R** definieren, dann ist unmittelbar klar, daß der Junktor $\to$ in **R** keine objektsprachliche Wiedergabe der Relation $\vdash$ sein kann. Damit das der Fall wäre, müßte die *Deduktionsäquivalenz*

(Ded) $$X, A \vdash B \text{ gdw } X \vdash A \to B$$

für Ableitbarkeit in **R** gelten. Aber in der $\to$-einführenden Richtung von links nach rechts gilt sie nicht. Man betrachte nur den Fall $X = \{B\}$ und führe den Pfeil zweimal auf der rechten Seite ein. Dann wäre die *Quodlibet*-Formel $B \to (A \to B)$ abgeleitet!

Wir können das so zusammenfassen: Wenn der Pfeil $\to$ in **R** ein relevant folgerndes Wenn-Dann objektsprachlich wiedergibt, dann kann die standardmäßig definierte Ableitbarkeitsrelation $\vdash$ keine solche Folgerungsrelation sein – selbst dann nicht, wenn sie auf **R** basiert. Aus der Perspektive der Relevanzlogik steht daher überhaupt nicht zu erwarten, daß die Beziehung zwischen $\to$ und $\vdash$ so ist, wie sie die Deduktionsäquivalenz behauptet.

Über den richtigen Gebrauch von Annahmen. Die letzte Beobachtung eröffnet die Suche nach einer Ableitbarkeitsrelation $\Vdash$, welche mit dem relevanten Pfeil im Sinne der Deduktionsäquivalenz harmoniert. Naheliegenderweise fragen wir zunächst, ob die Standardrelation $\vdash$ der Ableitbarkeit nicht so unter zusätzliche Bedingungen gestellt werden kann, daß eine Konklusion aus Annahmen genau dann ableitbar ist, wenn sie in einem *relevanzlogisch richtigen* Sinne aus diesen Annahmen folgt. Insbesondere würden wir in diesem Fall erwarten, daß $A \Vdash B$ genau dann der Fall ist, wenn $A \to B$ ein relevanzlogisches Theorem ist.

Um dieser Frage in möglichst einfacher Weise nachzugehen, betrachten wir zunächst das implikative Fragment $\mathbf{R}_{\to}$ von $\mathbf{R}$. In unserer Axiomatik wird $\mathbf{R}_{\to}$ festgelegt durch die reinen Implikationsaxiome Id, B, C, W, sowie die Regel MP. Auf dieser Basis sei eine *relevante Ableitung* eine Ableitung, welche die disjunktive Bedingung oben (1 oder 2 oder 3) *und* die folgende erfüllt:

4. Alle Annahmen in X werden für die terminale Formel A_n *gebraucht*.

Das genaue Verständnis dieser Bedingung wirft natürlich Fragen auf. Eine Ableitung aus Annahmen ist eine Folge von Formeln, die unter bestimmten Fortsetzungsregeln, nämlich (1-3), steht. Nach der Regel 2 dürfen wir jederzeit eine bisher gelungene Ableitung aus X mit einem Element aus X fortsetzen. Im Prinzip kann also jedes Element $A_i \in X$ beliebig oft in einer Ableitung aus X vorkommen. Bezieht sich nun der Gebrauch einer Formel auf ihr Vorkommen in der *Menge* X oder auf ihre möglicherweise multiplen Vorkommen in der Ableitungs*folge*? Um den Unterschied zu illustrieren, betrachten wir die Annahmenmenge $\{A\}$. Die Folge $\langle A \rangle$ ist eine Ableitung von A aus A in beiderlei Sinne; also $A \Vdash A$. Wie steht es um $\langle A, A \rangle$? Stellt $\langle A, A \rangle$ eine gute Ableitung ihres terminalen Elements aus der Menge $\{A\}$ dar? Nun, die (einzige) Annahme in der Menge $\{A\}$ wurde gebraucht, aber das erste ihrer zwei Vorkommen in der Folge $\langle A, A \rangle$ wurde nicht gebraucht. Den Unterschied können wir auch so fassen: Wenn wir unter einer Ansammlung von Annahmen eine Menge verstehen, dann macht es keinen Unterschied ob wir aus $\{A\}$ oder aus $\{A, A\}$ ableiten. $\{A\} \Vdash A$ stellt dann – im Sinn der Bedingung 4 – eine genausogute Folgerung dar wie $\{A, A\} \Vdash A$. Wenn aber 4 zwischen den beiden Folgerungen unterscheiden soll, dann muß (A) eine andere Annahmensammlung als (A, A) sein: eine, bei der es auf die Anzahl der Vorkommen derselben Formel ankommt. Wenn wir von (A) auf A schließen, gebrauchen wir alle Annahmen; wenn wir von (A, A) auf A schließen, tun wir das nicht. Wir werden gleich sehen, daß die Bedingung 4 so verstanden werden muß, daß zwischen $A \Vdash A$ und $A, A \Vdash A$ unterschieden werden kann. Also werden wir Annahmensammlungen nicht als Menge, sondern als *Multimengen* auffassen. (Wie in Mengen können in Multimengen Elemente permutiert werden; wie in Folgen kommt es jedoch auf die Anzahl der Vorkommen der Elemente an.)

Nach dieser Vorklärung können wir der Auswirkung der Bedingung 4 auf die drei Möglichkeiten in der Standarddefinition die folgende operationale Lesart geben:

> Jedes Vorkommen einer Annahme in der Multimenge X erhält einen sie identifizierenden Index i. Sodann erhält jede Formel A, die in eine Ableitung σ aus X eintritt, eine Indexmenge $ind(A)$ wie folgt:
>
> (a) Tritt ein *Axiom* A in σ ein, dann ist $ind(A) = \emptyset$.
> (b) Tritt eine *Annahme* A aus X mit Index i in σ ein, dann ist $ind(A) = \{i\}$.
> (c) Tritt B in σ als *Konklusion von MP* aus vorangehenden A und $A \to B$ ein, dann ist $ind(B) = ind(A) \cup ind(B)$.

Die drei Bedingungen versorgen also jede Formel, die in eine Ableitung eintreten kann, mit einer Indexmenge. Die Indexmenge der abzuleitenden Formel A_n soll nun anzeigen, daß alle Vorkommen von Formeln in X im Laufe der Ableitung *per* MP einen Beitrag geleistet haben. Das tut diese Menge genau dann, wenn $ind(A_n)$ aus genau den Annahmenindizes besteht.

Es gilt nun also $X \Vdash A$ genau dann, wenn es eine Folge $\langle A_0, ..., A_n \rangle$ gibt mit $A = A_n$. Jede der Formeln ist, wie gerade beschrieben, mit einer Indexmenge versehen. Darüberhinaus muß A nun diese Bedingung erfüllen:

4*. $ind(A) = \bigcup\{ind(B) : B \in X\}$.

Satz 32. Deduktionstheorem für $\mathbf{R}_{\to}$ (Moh 1950, Church 1951).
Wenn $A_1, ..., A_n, A \Vdash B$, dann $A_1, ..., A_n \Vdash A \to B$ ($n \geq 0$).

(Im Fall $n = 0$ lesen wir: Wenn $A \Vdash B$, dann $\Vdash A \to B$.) Es gilt natürlich auch die Umkehrung, sowie $\Vdash \subseteq \vdash$. Der Beweis verfährt im Prinzip wie für das Herbrand'sche Deduktionstheorem; nur müssen wir jetzt zusätzlich die Einhaltung der Bedingung 4 (in der Form 4*) überprüfen. Siehe [68].

Warum haben wir uns dafür entschieden Multimengen und nicht Mengen von Prämissen zu indizieren? Die Folge $\langle A^{\{i\}} \rangle$ ist eine Ableitung von A aus der Prämissensammlung (A^i), welche offensichtlich die Bedingung 4* erfüllt. Also haben wir $A \Vdash A$, woraus nach dem Satz folgt, daß $A \to A$ in $\mathbf{R}_{\to}$ ist. Wenn jedoch Prämissen in Mengen gesammelt werden, dann ist $(A^i) = (A^i, A^i)$, und also belegt $\langle A^{\{i\}} \rangle$ ebenso die Ableitbarkeitsbehauptung $A, A \Vdash A$, aus welcher nach dem Satz (zweimal) folgen würde, daß

Mingle $\qquad A \to (A \to A)$

ein Theorem von $\mathbf{R}_{\rightarrow}$ ist. Das ist aber nicht der Fall!

Nun ist Mingle zwar eine Instanz des *Quodlibet*-Schemas $A \rightarrow (B \rightarrow A)$, aber dieser Umstand allein kann Mingle sicher nicht relevanzlogisch disqualifizieren. Wäre es nicht besser, sich den Aufwand mit Multimengen zu sparen und Mingle an Bord zu nehmen? Leider entfaltet Mingle eine ungute Wirkung, wenn wir die üblichen und relevanzlogisch nicht beanstandeten Prinzipien für Negation, Konjunktion und Disjunktion zu $\mathbf{R}_{\rightarrow}$ hinzunehmen. Man betrachte die Mingle-Instanz

$$\neg A \rightarrow (\neg A \rightarrow \neg A).$$

Wenn wir das Konsequens kontraponieren und die Antezedenzien danach permutieren, erhalten wir das Schema

$$A \rightarrow (\neg A \rightarrow A).$$

Nun folgt eine Beobachtung von Meyer (in [10, §29.5]): Aus der Instanz

$$(A \rightarrow A) \wedge (B \rightarrow B) \rightarrow (\neg((A \rightarrow A) \wedge (B \rightarrow B)) \rightarrow (A \rightarrow A) \wedge (B \rightarrow B))$$

folgt nach den Regeln und Axiomen von $\mathbf{R}$

$$\neg(A \rightarrow A) \rightarrow (B \rightarrow B),$$

ein Schema mit Instanzen, welche die Bedingung der Variablenüberschneidung nicht erfüllen. Obgleich das implikative Fragment $\mathbf{RM}_{\rightarrow}$ ($\mathbf{R}_{\rightarrow}$ mit Mingle) gar nicht so schlecht aussieht, ist die Logik $\mathbf{RM}$ im angestrebten vollen Vokabular keine Relevanzlogik! Die Verwendung von Multimengen statt Mengen in der obigen Gebrauchsbedingung ist also – relevanzlogisch gesehen – die richtige Entscheidung.

So weit, so gut. Aber die Beschränkung auf Implikationsformeln ist ja nur ein Mittel, das Auffinden der richtigen Gebrauchsbedingungen zu erleichtern. Was immer wir finden, muß sich schließlich auch im vollen Vokabular bewähren. Wie steht es also um das Deduktionstheorem 32, wenn wir $\mathbf{R}_{\rightarrow}$ um naheliegende Axiome bzw. Regeln für weitere Junktoren erweitern? Die Präsenz der Konjunktionsoperation zwingt uns beispielsweise zur Adjunktionsregel

ADJ. $$\frac{A \quad B}{A \wedge B}.$$

Nun wäre es völlig natürlich zu verlangen, daß A und B ihre Indizes an $A \wedge B$ weitergeben – schließlich werden die beiden Annahmen für die Konjunktion

$A \wedge B$ in keinem anderen Sinne gebraucht, als A und $A \to B$ für B gebraucht werden. Aber dann können wir diese Ableitung konstruieren:

$$\dfrac{\dfrac{A^{\{1\}} \quad B^{\{2\}}}{A \wedge B^{\{1,2\}}}\,\text{ADJ} \qquad \dfrac{}{A \wedge B \to A^{\emptyset}}}{A^{\{1,2\}}}\,\text{MP}$$

($A \wedge B \to A$ ist ein Theorem und erhält daher – wie Axiome – die leere Indexmenge.) Wir haben A aus (A, B) abgeleitet unter Einhaltung der Bedingung 4*, d.h. $A, B \Vdash A$. Nach zweimaliger Anwendung des Deduktionstheorems erhalten wir $\Vdash A \to (B \to A)$! Die Adjunktionsregel stellt uns also vor ein Problem, dessen Lösung zu einer komplizierteren Antwort auf die Frage "Was heißt es, eine Prämisse zu gebrauchen?" zwingt. Die Antwort ist nun offensichtlich diktiert von der Absicht, die Logik **R** hervorzubringen. Der Begriff einer "guten" Ableitung motiviert nicht mehr die angestrebte Relevanzlogik, sondern umgekehrt: die vorgegebene Logik erzwingt kaum mehr als technische Bedingungen für Ableitungen, um die Deduktionsäquivalenz zu garantieren. So kann die Rede von "wirklich gebrauchten" Prämissen die Entscheidung für das System **R** nicht mehr unabhängig stützen.

An diesem Punkte gabelt sich der Weg. Relevanzlogiker wie Routley [254, §3.5] halten den Aspekt des Prämissengebrauchs für unwesentlich: In der Geschichte der Relevanzlogik spiele er eine gewisse heuristische Rolle, aus systematischer Sicht erweise er sich jedoch als letztlich nicht fruchtbar. Aus dieser Perspektive gibt es keinen starken Grund, daran festzuhalten, daß ein relevanzlogisches Wenn-Dann einen auf Prämissengebrauch basierenden Ableitungsbegriff spiegeln muß. Wir sehen den Grund, warum die Deduktionsäquivalenz für relevante Implikation problematisch ist – genauso, wie wir das in der Modallogik sehen. Diese Problematiken werfen auf die Postulate für $\to$ genausowenig einen Schatten wie auf die für $\Box$ – wenn man die Sache nur richtig versteht.

Dagegen halten Logiker wie Makinson [203, 204, 205] gerade die Idee, Schlüsse auszuzeichnen, die für eine Konklusion auf Prämissenresourcen tatsächlich zugreifen und vollständig verbrauchen, für dasjenige Leitmotiv, welches Relevanzlogik überhaupt interessant macht. Der natürliche und schon sehr früh verfolgte Ansatz, nämlich die Betrachtung des Haushaltens mit Prämissen in Systemen Natürlichen Schließens, scheint verbraucht und – wie wir gesehen haben – nicht weit genug zu führen. Makinson öffnet daher eine neue Perspektive, indem er Widerlegungen in Bäumen betrachtet und danach fragt, ob alle Information, welche eine zu prüfende Formel zur Verfügung stellt, tatsächlich für eine Widerlegung genutzt werden muß. Auch diese Perspektive zeichnet jedoch nicht eines der Standardsysteme der Relevanzlogik aus, sondern mündet natürlich eher in Systeme unter Einschluß des Mingle-Schemas $A \to (A \to A)$.

Über die richtige Bündelung von Annahmen. Abschließend wollen wir einen etwas anderen Versuch betrachten, die im Deduktionstheorem festgehalte Harmonie zwischen der Ableitbarkeitsbeziehung und der Implikation für Relevanzlogiken herzustellen. Anders ist dieser Versuch, da er das Augenmerk auf die Struktur von Annahmen richtet.

Der Ansatz richtet zugleich den Blick auf einen anderen wichtigen Theoriestrang in der Entwicklung der Relevanzlogik: die Beweistheorie im Sinne von Gentzens Sequenzenkalkülen (Gentzen [106]: "Schlußweisenkalküle"). In diesen Kalkülen ist der infrage stehende Teil der Deduktionsäquivalenz (Ded) einfach die Regel der Pfeil-Einführung. Für den Relevanzlogiker besteht die Aufgabe also nicht darin, (Ded) als eine Beobachtung über das Kalkül herzuleiten, sondern darin, ein relevanzlogisches Sequenzenkalkül mit der primitiven Einführungsregel für den Pfeil einzurichten – gewissermaßen also, die "Bedingung der Möglichkeit" einer solchen Regel zu schaffen.

In der Relevanzlogik gibt es schon früh eine ganze Reihe von Versuchen, Sequenzenkalküle zu entwickeln. Der jetzt vorzustellende Ansatz geht zurück auf Arbeiten von Belnap und Dunn, und wurde von Read [240] und Slaney [267] entwickelt. [28] Die Grundidee ist sehr einfach und fußt auf der Beobachtung (vgl. p. 383), daß es in Relevanzlogiken (im Prinzip) zwei Arten von "Konjunktionen" gibt: eine extensionale Boole'sche ($\wedge$) und eine intensionale relevante ($\circ$, Fusion). Analog zur Boole'schen Gleichung $A \wedge B = \neg(A \supset \neg B)$ können wir, falls gewünscht, die Fusion $A \circ B$ als kurz für $\neg(A \to \neg B)$ einführen. Nun gilt in Relevanzlogiken zwar

(1) $$\text{wenn } \vdash A \circ B \to C, \text{ dann } \vdash A \to (B \to C),$$

jedoch offensichtlich (sei $C = A$!) nicht

(2) $$\text{wenn } \vdash A \wedge B \to C, \text{ dann } \vdash A \to (B \to C).$$

Die extensionale Aggregation des Antezedens in (2) ist nicht von der richtigen Art, um daraus ein relevantes Antezedens für den Pfeil zu gewinnen. Dagegen unterstützt die intensionale Verknüpfung von A und B in (1) die Überführung in eine Implikation. Ferner ist natürlich $A \wedge B \to A$ ein relevanzlogisches Theorem; aber ebenso natürlich kann $A \circ B \to A$ keines sein (siehe (1)). Konjunktion $\wedge$ und Fusion $\circ$ sind also zwei grundlegend

[28] Meyers [212] deutet die Idee schon auf interessante Weise an.

verschiedene Arten, Information im Wenn-Teil einer Implikation zu aggregieren. Damit Ableitbarkeits- mit Wenn-Dann-Behauptungen harmonieren können, sollte sich diese wesentliche Unterscheidung auf der linken Seite von $\rightarrow$ auch auf der linken Seite der Ableitbarkeitsrelation (hier besser: der Sequenzenrelation) spiegeln können. So wie wir die Sache bisher betrachtet haben, tut sie das aber nicht: Wir haben bisher Annahmen immer nur als Mengen gebündelt. Kein Wunder also, daß relevante Implikation und klassische Ableitbarkeit nicht im Sinne der Deduktionsäquivalenz zueinander passen.

Um dieses Mißverhältnis zu beheben, wollen wir nun Annahmen auf zweierlei Weise bündeln: zum einen, in einer Weise, die der extensionalen Konjunktion entspricht, und zum anderen so, wie das in einer Formel durch die intensionale Fusion dargestellt wird. Die extensionale Bündelung von Annahmen sei durch ein Komma, die intensionale durch ein Semikolon bezeichnet.

Wir beginnen mit der Definition eines *Annahmenbündels* (kurz: Bündel).

1. Jede Formel sei ein Bündel.
2. Wenn X und Y Bündel sind, dann seien auch X, Y und $X; Y$ Bündel.

Nichts sonst ist ein Bündel. Wenn X ein Bündel ist, in dem das Bündel Y vorkommt, dann deuten wir das so an: $X(Y)$. Ein im selben Kontext folgendes $X(Y')$ deute an, daß in X das Teilbündel Y durch Y' überall ersetzt wurde.

Im nächsten Schritt charakterisieren wir die Bündelungsoperation mithilfe einer Relation $\leq$, die den Informationsgehalt von Bündeln vergleicht. Zunächst das Komma:

KB. $X, (Y, Z) \leq (X, Y), Z.$
KC. $X, Y \leq Y, X.$
KW. $X, X \leq X.$
KK. $X \leq X, Y.$

Intuitiv gesprochen, ist der Vergleich so angelegt, daß wenn X in Y enthalten ($\leq$) ist, dann soll auch aus Y folgen, was aus X schon folgt. Aus KW und KK ergibt sich unmittelbar, daß $X = X, X$. Über das Semikolon wollen wir bis auf weiteres nichts annehmen.

Nun definieren wir eine Folgerungsrelation $\Vdash$ (zwischen Bündeln und Formeln) so, wie es für Sequenzenkalküle typisch ist. Den Anfang bildet die Grundsequenz (auch "Annahmenregel" genannt):

Ann. $$A \Vdash A\ .$$

Die strukturellen Regeln können wir in einer einzigen generischen Regel zusammenfassen:

$$(\leq) \qquad \frac{X(Y) \Vdash A}{X(Y') \Vdash A} \text{ falls } Y \leq Y'.$$

Schließlich folgen die Einführungs- und Beseitigungsregeln für die Junktoren:

$$\frac{X \Vdash A \quad Y \Vdash B}{X, Y \Vdash A \wedge B}\wedge\text{Ein} \qquad \frac{X \Vdash A \wedge B}{X \Vdash A} \quad \frac{X \Vdash A \wedge B}{X \Vdash B}\wedge\text{Aus}$$

$$\frac{X \Vdash A}{X \Vdash A \vee B} \quad \frac{X \Vdash B}{X \Vdash A \vee B}\vee\text{Ein} \qquad \frac{X \Vdash A \vee B \quad Y(A) \Vdash C \quad Y(B) \Vdash C}{Y(X) \Vdash C}\vee\text{Aus}$$

$$\frac{X; A \Vdash B}{X \Vdash A \to B}\to\text{Ein} \qquad \frac{X \Vdash A \quad Y \Vdash A \to B}{X; Y \Vdash B}\to\text{Aus}$$

$$\frac{X; A \Vdash \neg B \quad Y \Vdash B}{X; Y \Vdash \neg A}\neg\text{Ein} \qquad \frac{X \Vdash \neg\neg A}{X \Vdash A}\neg\text{Aus}$$

Die Frage, warum in ¬Ein die intensionale Prämissenbündelung gebraucht wird, beantwortet sich mit der Beobachtung, daß diese (Reductio-) Regel äquivalent ist (gegeben →Ein) zur Modus Tollens-Regel

$$\frac{X \Vdash A \to B \quad Y \Vdash \neg B}{X; Y \Vdash \neg A}.$$

Offenbar fehlt hier noch die Einführung der doppelten Negation,

$$\frac{X \Vdash A}{X \Vdash \neg\neg A}\text{DNI}.$$

Diese bekommen wir gratis, wenn wir das Vokabular um die ohnehin nützliche intensionale Verum-Konstante t erweitern, mit der Semikolonregel

St. $\quad t; X = X.$

Die t-Konstante ist also ein idempotentes, "leeres" Bündel für die Semikolonoperation. Jetzt können wir DN so ableiten:

$$\dfrac{\dfrac{X \Vdash A \qquad \dfrac{\neg A \Vdash \neg A}{t; \neg A \Vdash \neg A}\text{St}}{t; X \Vdash \neg\neg A}\neg\text{Ein}}{X \Vdash \neg\neg A}\text{St}$$

Mit dieser kurzen Ableitung haben wir schon die übliche Definition einer beweisbaren Sequenz illustriert:

> Eine *Sequenz* $X \Vdash A$ ist genau dann *beweisbar* (im beschriebenen Sequenzenkalkül), wenn es eine Folge von Sequenzen gibt mit $X \Vdash A$ als terminalem Element und so, daß jedes Element der Folge entweder eine Annahmensequenz ist, oder aus vorhergehenden Elementen aufgrund einer der strukturellen oder operationalen Regeln hervorgeht.
>
> Wir nennen eine *Formel* A genau dann *beweisbar*, wenn die Sequenz $t \Vdash A$ beweisbar ist.

Da es bisher – außer St – keine strukturellen Regeln gibt, die es erlauben, aus Semikolon-Sequenzen neue Sequenzen abzuleiten, definiert dieses Kalkül eine sehr schwache Logik, welche noch unterhalb der Modellierungsschwelle der in diesem Kapitel behandelten Rahmentypen bleibt. Die Menge der beweisbaren Formeln koinzidiert mit den Theoremen einer als **DW** bekannten Logik.[29] Eine Axiomatisierung von **DW** entsteht, wenn man aus den Axiomen für **R** (siehe p. 395) die Axiome B, C, und W streicht und eine neue (in **R** ableitbare) Regel hinzunimmt:

$$\frac{A \to B \qquad C \to D}{(B \to C) \to (A \to D).}$$

Damit ist der Weg vorgezeichnet, wie man **DW** zur Stärke von **R** erweitern kann: Wir müssen Semikolonregeln angeben, welche den Axiomen B (Präfix), C (Permutation) und W (Kontraktion) entsprechen:

SB. $X;(Y;Z) \le (X;Y);Z.$

SC. $X;Y \le Y;X.$

SW. $X;X \le X.$

Wenn wir SB und SC hinzunehmen, dann sind die im Sequenzenkalkül beweisbaren Formeln genau die Theoreme von **C**. Nehmen wir noch SW hinzu, so erhalten wir **R**. Die Regel

SK. $X \le X;Y$

kollabiert die Unterscheidung zwischen Komma und Semikolon – was uns zurück zur klassischen Logik bringt.

Sequenzenkalküle axiomatisieren Folgerungsrelationen. Abstrakt, d.h. ohne Rücksicht auf die Junktoren einer Sprache, werden diese allein durch strukturelle sowie die Annahmenregel charakterisiert. Diese allgemeine

[29] Die Benennung geht auf Brady zurück; siehe [36, 37].

Theorie des Folgerns bildet einen Kontext, in den die Junktorenregeln eingebettet werden. Wir können hier eine Analogie bilden zur Semantik indexikalischer Ausdrücke nach Kaplan [160]. Ein Ausdruck wie "hier" hat einen konstanten *Charakter*, der je nach Kontext einen spezifischen *Inhalt* (*content*) annimmt. Im übertragenen Sinne wird der Charakter eines Junktors durch ein Paar von Junktorenregeln angegeben. Dieser Charakter legt die Bedeutung eines Junktors – seinen Inhalt – noch nicht fest, sondern tut dies erst, wenn er in eine bestimmte Theorie des Folgern eingebettet wird.

In dem Ansatz, den wir gerade vorgestellt haben, folgt die Deduktionsäquivalenz

$$X; A \Vdash B \text{ gdw } X \Vdash A \to B$$

unmittelbar aus dem Charakter der Implikation (sowie der Grundsequenz $A \Vdash A$). In diesem Sinne drückt die Äquivalenz die gewünschte *begriffliche* Beziehung zwischen Folgerung und Implikation aus.

Der Ansatz führt vor, wie sich der Inhalt der Junktoren im Zusammenspiel mit den strukturellen Regeln ändert. Dabei gibt es einen einfachen Weg von der Basislogik zum System **R** und von dort zur klassischen Logik **KL**. Über den Charakter der Junktoren gibt es zwischen relevanter und klassischer Logik keinen Dissens. Die Differenz zwischen **R** und **KL** liegt allein in der allgemeinen Theorie der Folgerung und wird mit der Regel SK klar benannt. Diese Regel betrifft nur den Inhalt der $\to$-Regeln und der Reductio-Regel $\neg$-Ein – die ja äquivalent zu einer Modus Tollens-Regel ist. So wird auf einfache Weise deutlich, wie **R** durch einen minimalen Eingriff in **KL** die *Quodlibet*-Sequenzen ausschließt.

10. Konkurrenz oder Ergänzung?

Wie intuitionistische Logik ist Relevanzlogik ein zwiespältiges Unternehmen. Im Gegensatz zur Modallogik werden beide meist als nichtklassische Logiken bezeichnet. Damit wird in den Vordergrund gestellt, daß beide Logiken als Gegenentwurf zur klassischen Logik verstanden werden können und von vielen ihrer Proponenten auch genau so verstanden werden. Sie werden als Antworten auf die Frage nach der "richtigen" Negation bzw. Implikation (oder beiden) präsentiert. Aber allein die Frage ist schon problematisch. Eine logische Frage kann das kaum sein, denn welche logischen Methoden könnten sie beantworten? Ist es also eine metaphysische oder linguistische Frage?

Der Gegensatz zur Modallogik ist schon aus zwei Gründen wenig zwingend. Erstens, kann man auch in der Modallogik eine dezidiert anti-klassische Theorie sehen. In der Anfangsphase der Modallogik war das die Auffassung von C.I. Lewis, der in der strikten Implikation die “bessere” Implikation sah; siehe z.B. [178]. Für Quine war die Modallogik bekanntermaßen eine illegitime Abkehr vom Extensionalismus der klassischen Logik. Zweitens – und hier wichtiger – läßt sich auch die Relevanzlogik als eine Erweiterung der klassischen Logik darstellen. Im wesentlichen [30] genügt es dazu, beispielsweise die Sprache des Systems **R** um die Boole'sche Negation $-$ zu erweitern, mit der Interpretationsbedingung

$$(-) \qquad a \models -A \text{ gdw } a \not\models A.$$

Mit $\{-, \wedge\}$ haben wir dann ein funktional vollständiges Vokabular unter der klassischen Interpretation, dem die Relevanzlogik die intensionalen Junktoren $\neg$ und $\rightarrow$ hinzufügt; siehe [212, 213, 214]. Je nach Ausgangsperspektive ist das Resultat eine konservative Erweiterung der klassischen Logik **KL** oder der Relevanzlogik **R**.

Es ist also sicher nicht zwingend, Relevanzlogik als Konkurrenz zur klassischen Logik aufzufassen. Wie ergänzt dann Relevanzlogik die klassische Logik? Im Prinzip nicht anders als die Modallogik. Die Modallogik ist die Theorie eines monadischen Satzoperators, der in Punkte-Modellen mit einer zweistelligen Relation in der bekannten Klausel ($\Box$) interpretiert wird. Die Relevanzlogik ist die Theorie eines dyadischen Satzoperators, den wir in solchen Modellen mit einer dreistelligen Operation in der Klausel ($\rightarrow$) interpretieren. Die angestrebte Interpretation ist im einen Fall paradigmatisch durch den Begriff einer irgendgearteten Notwendigkeit vorgegeben, im anderen Fall durch den Begriff eines Konditionals. Modallogik ist so eine Rahmentheorie, in der bestimmte Begriffe von Notwendigkeit – einschließlich epistemischer oder deontischer Varianten – logisch betrachtet werden können. Im selben Sinne bietet Relevanzlogik einen weiten logischen Rahmen an, in dem Konditionale verschiedenster Art studiert werden können. (Siehe dazu auch den Abschnitt auf pp. 375ff.) In beiden Rahmen gibt es Grenzfälle, insbesondere Notwendigkeit als Wahrheit bzw. Wenn-Dann als materiales Konditional.

Wir sehen nun, daß unsere Übertragung der Kaplan'schen Unterscheidung zwischen Charakter und Inhalt nicht ganz so metaphorisch ist, wie sie auf den ersten Blick ausgesehen haben mag. Notwendigkeit und Wenn-Dann sind Begriffe, die eine abstrakte Kernbedeutung – einen “Charakter”

[30] Die Rahmen müssen etwas nachjustiert werden. Meyer und Routley erklären in [214] warum und wie.

– haben, die durch eine konstante Interpretationsklausel gegeben ist. Erst wenn diese Kernbedeutung einen Rahmen zu einem Modell ergänzt, generiert sie Wahrheitsbedingungen ("Inhalt").[31] Die Relevanzlogik mag – wie die intuitionistische Logik – als ein Reformprogramm begonnen haben. Aber sie ist weit darüber hinausgewachsen zu einem Arsenal vielfältiger und sehr weitgehend integrierter Perspektiven auf die Logik vieler Arten von Konditionalaussagen.

[31] Die Analogie zu Kaplans Behandlung indexikalischer Ausdrücke ist natürlich nicht völlig strikt. Die hier behandelten Junktoren sind keine indexikalischen Ausdrücke.

VII.

ANFECHTBARES SCHLIESSEN

1. Formale und materiale Logik

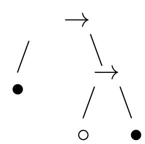
Von Logik wird gewöhnlich gesagt, sie sei “formal”. Was das genau heißt, ist am Ende gar nicht so einfach zu erklären. Aber die Idee ist recht einfach: Wenn eine Aussage den Status logischer Wahrheit genießt, dann tut es jede Aussage der gleichen Form ebenso. Logische Wahrheit ist unter gleichförmiger Einsetzung abgeschlossen. So ist die Aussage

$$(*) \qquad P \to (Q \to P)$$

mit Atomen P und Q einer Sprache $\mathcal{L}$, logisch wahr. Und gleichgültig wie komplex $\mathcal{L}$-Formeln A und B sein mögen, die wir für P und Q in $(*)$ einsetzen können, das Resultat der Einsetzung,

$$A \to (B \to A),$$

ist von derselben Form wie $(*)$, nämlich



und ebenfalls wahr. Dasselbe gilt analog für die Eigenschaft ein logisch gültiger Schluß zu sein.

Der Umstand, daß Logik formal ist, reflektiert die Vorstellung, daß die logischen Wahrheiten *unabhängig vom Gegenstand der Rede* wahr sind. Es kommt nicht darauf an, was A und B (im Beispiel) beschreiben – ja nicht einmal darauf, ob sie überhaupt etwas beschreiben –, sondern nur auf die Art und Weise, wie sie durch den Junktor $\rightarrow$ zusammengebunden sind: nur auf die Struktur der Formel, nicht auf ihren Inhalt.

In der Axiomatik der Logik wird die Bedingung, daß Logik formal sei, durch die spezielle Schlußregel der gleichförmigen Einsetzung oder durch die Angabe der Axiome in schematischer Form realisiert. D.h. entweder wird eine konkrete Formel als Axiom angegeben und die Ableitbarkeit gleichförmiger Formeln durch die Substitutionsregel gewährleistet, oder es werden alle Formeln einer bestimmten Form durch eine schematische Beschreibung als Axiome spezifiziert. Die beiden Verfahren sind ersichtlich äquivalent.

Auf der Ebene der Modelle wird Formalität durch Quantifikation über eine eingeschränkte ("zulässige") Klasse von Modellen garantiert. Wenn wir eine aussagenlogische Sprache betrachten, deren Formeln aus einer bestimmten Menge ATM von Atomen und Junktoren, z.B. $\neg$ und $\wedge$ generiert werden, dann ist ein klassisches aussagenlogisches Modell dieser Sprache eine Interpretation I, welche Atome in genau einen der zwei Werte 0 und 1 und die beiden Junktoren in Funktionen dieser Wert abbildet. Wenn wir nun logische Wahrheit definieren als Wahrheit in "allen" Modellen, dann betrachten wir nicht wirklich *alle* Modelle der Sprache. Während wir die Interpretation der Atome maximal variieren lassen, d.h. alle Abbildungen $I : \mathrm{ATM} \rightarrow 0, 1$ betrachten, wird die Interpretation von $\neg$ und $\wedge$ überhaupt nicht über der Menge der Wahrheitsfunktionen variiert. Die Junktoren $\neg$ und $\wedge$ werden vielmehr konstant in die Komplementärfunktion bzw. die Minimum-Funktion abgebildet.[1] Deshalb nennt man die Junktoren auch *logische Konstanten*. Ein zulässiges klassisches Modell der betrachteten Sprache ist also eines, in dem, erstens, I die sehr weite Bedingung erfüllt, Atome in die Werte 0 und 1 abzubilden, und zweitens, I die sehr enge Bedingung erfüllt, $\neg$ und $\wedge$ in die gerade genannten *bestimmten* Wahrheitsfunktionen abzubilden. Logische Wahrheit ist dann Wahrheit in der Klasse aller zulässigen Modelle.

Die Klasse der zulässigen Modelle könnte man im Prinzip auf zweierlei Weise erweitern. Die interessantere Option ist es, die Interpretation, d.h. den Wahrheitswert bestimmter Atome oder Zusammensetzungen von Atomen konstant zu halten. In diesem Fall behandeln wir bestimmte Formeln genau wie logische Konstanten. Freilich definiert dann die Quantifika-

[1] Der Wert von $\neg A$ ist 1 vermindert um den Wert von A; der Wert von $A \wedge B$ ist der kleinere der Werte von A und von B.

tion über eine so erweiterte Klasse von Modellen nicht logische Wahrheit schlechthin, sondern logische Wahrheit unter materialen Annahmen. Die Quantifikation über alle Modelle, in denen die Annahme gilt, gibt Auskunft darüber, was nun logisch folgt – wir haben es sozusagen mit einer *materialen Logik* zu tun.

Um es einfach und konkret zu veranschaulichen: Sei P eine atomare Formel und sei $\mathbb{K}^P$ die Klasse aller (klassischen) Modelle, d.h. Interpretationen $I : \mathrm{ATM} \longrightarrow \{0, 1\}$, mit $I(P) = 1$. Ferner sei $\mathbf{KL}^P$ die Menge aller Formeln, die in allen Modellen in $\mathbb{K}^P$ wahr sind.

Zwei Eigenschaften der "materialen Logik" $\mathbf{KL}^P$ sind evident:

1. $\mathbf{KL}^P$ ist eine echte Erweiterung der klassischen Logik $\mathbf{KL}$: $\mathbf{KL}^P \subset \mathbf{KL}$. In diesem Sinne ist $\mathbf{KL}^P$ *stärker* als $\mathbf{KL}$.

2. $\mathbf{KL}^P$ ist nicht abgeschlossen unter gleichförmiger Einsetzung. Denn das würde hier bedeuten, daß $\mathbf{KL}^P = \mathbf{KL}^A$ für beliebige Formeln A, was sicher nicht der Fall ist. $\mathbf{KL}^P$ ist also keine *formale* Logik in dem Sinne, in dem wir das oben erläutert haben.

Da $\mathbf{KL}^P$ gegenüber $\mathbf{KL}$ nicht unter der Einsetzungsregel abgeschlossen ist, sagt man manchmal auch $\mathbf{KL}^P$ sei "schwächer" als $\mathbf{KL}$. (Und im selben Sinne sagt man manchmal, nichtmonotone Logiken seien schwächer als die (monotone) klassische Logik.) Aber diese Redeweise wollen wir uns hier nicht zu eigen machen.

Alle Logiken, mit denen wir es in diesem Kapitel zu tun haben werden, sind nicht-formale (materiale) Erweiterungen der klassischen Logik. Für formale echte Erweiterungen der klassischen Aussagenlogik gibt es nur einen Grenzfall: die Menge aller Formeln. Bevor wir uns davon überzeugen, wollen wir zunächst unsere Aufmersamkeit von (logisch) wahren Formeln zu (logisch) gültigen Folgerungen verlagern, denn es ist diese Relation, die für dieses Kapitel grundlegend sein wird.

Eine Folgerungsrelation $\vdash$ setzt Formelmengen einerseits zu Formeln andererseits in Beziehung, d.h. $\vdash \subseteq \wp(\mathrm{FML}) \times \mathrm{FML}$. Die Elemente der Menge nennen wir die *Prämissen*, die Formel nennen wir die *Konklusion* der Folgerung. Anders als in vorhergehenden Kapiteln steht $\vdash$ in diesem Kapitel nicht immer für eine rein syntaktisch definierte Relation. Manchmal werden wir $\vdash$ im Hinblick auf Modelle definieren – wofür wir in anderen Kapiteln die Notation $\models$ bevorzugt hätten. Manchmal wechseln wir zwanglos zwischen syntaktisch und semantisch definierbaren Eigenschaften von $\vdash$. Die Folgerungsrelation $\vdash$ wird immer wechselseitig definierbar sein mit einer Folgerungsoperation Cn; davon gleich mehr.

Hier soll nun $\vdash$ die klassische Folgerungsrelation sein, d.h.

$$X \vdash A \text{ gdw } \forall I: \text{ wenn } I(X) = 1, \text{ dann } I(A) = 1,$$

wobei $I(X) = 1$ kurz ist für: $I(B) = 1$, für alle Formeln B in der Formelmenge X.

Was wir gerade über $\mathbf{KL}^P$ mit Atom P gesagt haben, läßt sich zu beliebigen Erweiterungen $\mathbf{KL}^K$ mit Formelmenge K verallgemeinern und auf die "materiale" Folgerungsrelation $\vdash_K$ übertragen, wenn wir diese so verstehen:

$$X \vdash_K A \text{ gdw } \forall I \text{ mit } I(K) = 1: \text{ wenn } I(X) = 1, \text{ dann } I(A) = 1.$$

1. Die Relation $\vdash_K$ ist eine echte Erweiterung der klassischen Relation logischer Folgerung $\vdash$. In diesem Sinne ist $\vdash_K$ *stärker* als $\vdash$.
2. Die Relation $\vdash_K$ ist im Gegensatz zu $\vdash$ nicht abgeschlossen unter gleichförmiger Einsetzung. Sie differenziert zwischen Prämissen und Konklusionen gleicher Form: Einige Prämissen-Konklusions-Paare stehen in der Relation der Folgerung zueinander, während andere Paare gleicher Form es nicht tun. In diesem Sinne ist $\vdash_K$ keine *formale* Folgerungsrelation.

Der nächste Satz hält die Beobachtung fest, daß die klassische Folgerungsrelation deduktiv (oder Post-) vollständig ist: Es gibt nur *eine* formale, d.h unter gleichförmiger Einsetzung abgeschlossene Erweiterung von $\vdash$, nämlich die triviale Relation $\wp(\mathrm{FML}) \times \mathrm{FML}$. Um den Satz zu beweisen, erweitern wir zunächst die Definition einer gleichförmigen Einsetzungsfunktion σ für Formeln zu einer Funktion σ^* für Paare der Form (X, A) mit $X = \{A_1, ..., A_n\} \subseteq \mathrm{FML}$ und $A \in \mathrm{FML}$:

$$\sigma^*(X, A) = (\{\sigma A_1, ...\sigma A_n\}, \sigma A).$$

Da diese Erweiterung völlig natürlich ist, schreiben wir im folgenden für die Basis- wie für die erweiterte Funktion einfach σ (und $\sigma(\{A_1, ..., A_n\})$ für $\{\sigma A_1, ..., \sigma A_n\}$). Sodann beweisen wir ein Lemma.

LEMMA 1. *Wenn $X \nvdash A$, dann gibt es eine Einsetzung $\tau(X, A)$ so, daß jede Formel in $\tau(X)$ äquivalent zu $\top$ und τA äquivalent zu $\bot$ ist.*

BEWEIS. Es sei I eine Interpretation und A eine Formel, in der ein Atom P vorkommt. Wenn $I(P) = 1$, dann sei $\tau(A[P]) = A[\top]$; anderenfalls sei $\tau(A[P]) = A[\bot]$, für jedes Atom P in A. (Mit anderen Worten, τ ersetzt in A alle wahren Atome durch $\top$ und alle falschen Atome durch $\bot$.) Eine einfache Induktion über den Aufbau von A zeigt:

$$(*) \qquad I(A) = 1 \text{ gdw } \tau A \equiv \top, \quad \text{und} \quad I(A) = 0 \text{ gdw } \tau A \equiv \bot.$$

Wir betrachten nun $\tau(X, A)$. Wenn $A \not\vdash x$, dann gibt es eine Interpretation I mit $I(X) = 1$ und $I(A) = 0$. Also haben wir nach $(*)$ $\tau A \equiv \bot$ und für jede Formel B in X, $\tau B \equiv \bot$. ∎

Satz 2. *(Post 1921) Es sei $\vdash^+$ eine echte Erweiterung der klassischen Folgerungsrelation $\vdash$. Wenn $\vdash^+$ unter gleichförmiger Einsetzung abgeschlossen ist, dann ist $\vdash^+$ trivial, d.h. $\vdash^+ = \wp(FML) \times FML$.*

Beweis. Wir nehmen an, daß $\vdash\ \subset\ \vdash^+$. Dann gibt es $X \subseteq$ FML und $A \in$ FML mit (1) $X \not\vdash A$ und (2) $X \vdash^+ A$. Da wir jetzt annehmen, $\vdash^+$ sei unter Einsetzung abgeschlossen, so folgt aus (2),

(3) für alle Einsetzungen σ: $\sigma X \vdash^+ \sigma A$.

Aus (1) folgt, daß es eine Interpretation I gibt so, daß $I(X) = 1$ und $I(A) = 0$. Auf der Basis von I definieren wir die Einsetzung τ wie im Lemma und haben

$$(4) \qquad \tau X \equiv \top \quad \text{und} \quad \tau A \equiv \bot.$$

Es folgt aus (3) und (4), daß (5) $\top \vdash^+ \bot$. Nun haben wir aber auch $Y \vdash \top$ und $\bot \vdash B$, für beliebige Y und B. Da $\vdash\ \subseteq\ \vdash^+$, gilt ebenso:

$$(6) \qquad Y \vdash^+ \top \quad \text{und} \quad \bot \vdash^+ B.$$

Aus (5) und (6) folgt nun nach der Transitivitätsregel (zweimal)

$$\frac{X \vdash^+ A \qquad A \vdash^+ B}{X \vdash^+ B},$$

daß $Y \vdash^+ B$. ∎

Daß die Transitivitätsregel bzw. allgemeiner die *Schnittregel*[2]

$$\textit{Schnitt} \qquad \frac{\forall A \in Y : X \vdash A \qquad X, Y \vdash B}{X \vdash B}$$

im Beweis auch für $\vdash^+$ gilt, macht man sich schnell klar anhand einer verallgemeinernden Betrachtung, die von unserer Definition von $\vdash_K$ (S. 414) ausgeht.

[2] Englisch: *Cut*, auch *Cumulative Transitivity*.

Anmerkung. Aussagen wie $\forall A \in Y : X \vdash A$ wollen wir künftig auch so abkürzen: $X \vdash \forall Y$. Danach sieht die Schnittregel so aus:

Schnitt $$\frac{X \vdash \forall Y \qquad X, Y \vdash B}{X \vdash B}$$

In der Literatur über nichtmonotone Logik wird die Schnittregel auch *kumulative Transitivität* genannt. Die Schnittregel wird manchmal als eine "verallgemeinerte" Transitivitätsregel (aus $X \vdash \forall Y$ und $Y \vdash B$ schließe $X \vdash B$) eingeführt. Aber Transitivität in diesem Sinne folgt aus Schnitt nur, wenn wir Monotonie (aus $Y \vdash B$ folgt $X, Y \vdash B$) voraussetzen, was wir in diesem Kapitel natürlich selten tun dürfen.

Materiale Folgerung ist, vom Standpunkt der logischen Theorie betrachtet, keine wesentlich neue Relation. Denn für beliebig gewählte Formelmengen K bedeutet $X \vdash_K A$ nichts anderes als $X, K \vdash A$. Interpretationen, die Formeln in K falsch machen, zu ignorieren, heißt einfach K den Status einer Annahme für Schlüsse im Sinne von $\vdash$ zuzuweisen. Wenn wir vor dem Hintergrund K auf A aus den Annahmen X schließen, dann schließen wir auf A aus den Annahmen $X \cup K$ (und umgekehrt).[3]

Dennoch können wir in Bezug auf materiale Folgerungsrelationen $\vdash_K$ eine Frage von der Art stellen, die für die Behandlung anfechtbaren Schließens in diesem Kapitel eine Schlüsselrolle spielen wird und die schon für diese sehr einfache und eigentlich nicht weiter interessante Folgerungsrelation auf eine interessante Antwort trifft. Wir fragen: Gegeben eine Relation $\vdash^+ \subseteq \wp(\mathrm{FML}) \times \mathrm{FML}$, gibt es rein syntaktische Bedingungen (etwa von der Art der Schnittregel), welche $\vdash^+$ genau dann erfüllt, wenn $\vdash^+$ sich als eine materiale Folgerungsrelation darstellen läßt, d.h. wenn es eine Menge K gibt so, daß $\vdash^+ = \vdash_K$? Eine Angabe von auf diese Weise charakteristischen Bedingungen für materiale Folgerungsrelationen nennt man ein *Repräsentationsresultat* für solche Relationen.

Im nächsten Satz werden neben der Schnittregel die folgenden Bedingungen für $\vdash^+$ genannt (X_f deutet eine *endliche* Menge an):

Supraklassisch $$\vdash \subseteq \vdash^+$$

Reflexiv $$X, A \vdash^+ A$$

Monoton $$\frac{X \vdash^+ A}{X, Y \vdash^+ A}$$

[3] Vgl. dazu auch die Überlegungen zu enthymematischem Schließen in Urchs' [284].

Kompakt Wenn $X \vdash^{+} A$, dann $\exists X_f \subseteq X : X_f \vdash^{+} A$

Disjunktiv $$\frac{X, A \vdash^{+} C \qquad X, B \vdash^{+} C}{X, A \vee B \vdash^{+} C}$$

Nur um den nächsten Satz kurz zu formulieren, wollen wir eine Folgerungsrelation $\vdash^{+}$ *monoton material* nennen, wenn sie alle diese Eigenschaften hat. Aus diesen Eigenschaften sind weitere ableitbar, die im Hinblick auf den folgenden Beweis einer Beobachtung wert sind ($X \equiv Y$ notiert die logische Äquivalenz von X und Y im Sinne der klassischen Relation $\vdash$; wir erweitern die Relation hier zwanglos von Formeln zu Formelmengen): [4]

Linksäquivalent $$\frac{X \equiv Y \qquad X \vdash^{+} A}{Y \vdash^{+} A}$$

Rechtsabschwächend $$\frac{X \vdash^{+} A \qquad A \vdash B}{X \vdash^{+} B}$$

SATZ 3. *Wenn $\vdash^{+}$ eine monoton materiale Folgerungsrelation ist, dann gibt es eine Formelmenge K so, daß $\vdash^{+} = \vdash_K$.*

BEWEIS. Sei $\vdash^{+}$ monoton material. Wir betrachten die Menge

$$K = \{A : \emptyset \vdash^{+} A\}$$

und zeigen:

1. Wenn $X \vdash^{+} A$, dann $K, X \vdash A$; und
2. wenn $K, X \vdash A$, dann $X \vdash^{+} A$.

Ad 1. Wir nehmen an, daß $X \vdash^{+} A$ und schließen nach Kompaktheit, daß es eine endliche Teilmenge X_f von X gibt mit $X_f \vdash^{+} A$. Es sei nun $\wedge X_f$ die Konjunktion aller Formeln in X_f; wir kürzen $\wedge X_f$ zu B ab. Da $X_f \equiv B$, so folgt aufgrund von Linksäquivalenz, daß $B \vdash^{+} A$, was wir rechts nach (1) $B \vdash^{+} A \vee \neg B$ abschwächen können. Da $\neg B \vdash A \vee \neg B$, so gilt auch (supraklassisch) (2) $\neg B \vdash^{+} A \vee \neg B$. Aus (1) und (2) folgt nun nach der Disjunkionsregel, $B \vee \neg B \vdash^{+} A \vee \neg B$, d.h. (Linksäquivalenz) $\emptyset \vdash^{+} A \vee \neg B$. Aufgrund der Definition von K wissen wir nun, daß $A \vee \neg B \in K$ und also, daß (3) $K \vdash A \vee \neg B$. Ebenso ist klar, daß (4) $X \vdash \wedge X_f$ $[= B]$. So folgt dann aus (3) und (4) (Disj. Syll.), daß $K, X \vdash A$, wie gewünscht.

Ad 2. Aus der Annahme $X, Y \vdash A$ und $\vdash \subseteq \vdash^{+}$ (supraklassisch) folgt, daß (1) $X, Y \vdash^{+} A$. Nach der Definition von X gilt aber $\emptyset \vdash^{+} \forall X$. Da $\vdash^{+}$ monoton ist, folgt weiter (2) $Y \vdash^{+} \forall X$ und also aus (1) und (2) durch Schnitt $Y \vdash^{+} A$. ∎

[4] Englisch: *Left Logical Equivalence* und *Right Weakening.*

Vorsichtige materiale Logik und Monotoniefehler. Es scheint hochriskant zu sein, manche kontingente Aussagen für genauso sicher wahr zu halten wie Tautologien, d.h. Aussagen wie $A \vee \neg A$. Tatsächlich tun wir aber im Alltag genau das, denn manche Aussagen sind für alle praktischen Zwecke ziemlich sicher wahr. Wir bewegen uns in einer weitgehend vertrauten Welt. Das heißt soviel wie: Auf die Wahrheit bestimmter Annahmen dürfen wir vertrauen. So vertrauen wir gewöhnlich auf die Annahme, daß in einem betankten und gut gewartetem Auto eine Drehung des Zündschlüssels den Motor startet – wir rechnen nicht damit, daß uns jemand einen Streich gespielt hat. Wer ein Zündholz und eine geeignete Reibefläche hat, der kann sicher eine Flamme erzeugen – müssen wir die Möglichkeit berücksichtigen, daß er es unter Wasser versucht? Solche und ähnlich entfernte – "an den Haaren herbeigezogene" – Möglichkeiten blenden wir im allgemeinen aus und halten so kontingente Annahmen fix – auf den ersten Blick so, als würden wir sie in die Nähe logischer Wahrheiten rücken. Täten wir das nicht, so würden wir unsere Umgebung als unberechenbar ansehen und unser Handeln wäre gelähmt oder ohne Zuversicht.

Funktionieren solche, als sicher unterstellten Hintergrundannahmen, wie die Formelmenge X einer monoton materialen Relation $\vdash_X$? Wenn wir etwas genauer hinsehen, dann erkennen wir, daß wir im Alltag nicht einfach mit monoton materialen Folgerungsrelationen arbeiten. Das wäre zu unvorsichtig und unvernünftig. Wir gehen für gewöhnlich nicht davon aus, daß nachts jemand auf der Straße an unserem Auto hantiert; vielmehr nehmen wir an, daß das nicht der Fall ist. Die Annahme zeigt sich darin, daß wir uns gewöhnlich einfach hinter das Steuer setzen und den Zündschlüssel drehen, statt zuvor unter die Motorhaube zu sehen. Aber obgleich wir annehmen, daß das Auto unmanipuliert ist, wäre es unvernünftig, die Möglichkeit der Manipulation *in jedem Fall* auszuschließen, d.h. gleichgültig, was wir über unser Auto erfahren sollten. Wir vertrauen der Annahme, aber nicht blind.

Das zeigt sich schon an einem trivialen Beispiel. Für gewöhnlich schließen wir so [5]

$$\textit{betankt},\ \textit{gewartet}\ \mid\!\sim \textit{startet}, \tag{1}$$

wobei zum Hintergrund dieses Schlusses auch die Annahme gehört, daß das Auto *unmanipuliert* ist. Aber sicher schließen wir *nicht* so:

$$\textit{betankt},\ \textit{gewartet},\ \textit{manipuliert}\ \mid\!\sim \textit{startet}. \tag{2}$$

[5] Hier und im folgenden benutzen wir das Zeichen $\mid\!\sim$ ("Schlange") für nichtmonotone Folgerungsrelationen: manchmal, wie hier, in generischer Weise, manchmal, wie in den folgenden Abschnitten, auf jeweils bestimmte Weise definiert.

Hier scheint die Annahme *manipuliert* eine oder mehrere der Hintergrundannahmen “auszuschalten”, die den ersten Schluß ermöglicht haben; dazu gehört die Annahme *unmanipuliert*. Das Beispiel zeigt sofort, daß die Relation $\mid\!\sim$ nicht unter Erweiterung der Annahmenmenge abgeschlossen ist, d.h. $\mid\!\sim$ ist eine *nichtmonotone* Folgerungsrelation. Im Gegensatz zu monoton materialen Relationen $\vdash^+$ handelt es sich bei $\mid\!\sim$ um eine *vorsichtig* materiale Folgerungsrelation. Wir lassen Vorsicht walten, indem wir die Anfechtbarkeit von Hintergrundannahmen (und Hintergrund*regeln*, wie wir später sehen werden) im Auge behalten. Nichtmonotones Schließen in genau diesem Sinne ist das Thema dieses Kapitels.

Wenn wir das Beispiel im Detail verstehen möchten, dann müssen wir die Hintergrundannahmen im Schluß (1) explizit machen und erklären, wie die zusätzliche Annahme *manipuliert* auf die Hintergrundannahmen so einwirkt, daß in (2) auf Konklusion *startet* nicht mehr geschlossen werden kann. Ganz allgemein wird jede Theorie vorsichtig materialen Schließens angeben müssen, auf welcher Grundlage gesetzte Prämissen Hintergrundannahmen ein- und ausschalten.

Um das Prinzip, nach dem wir hier nichtmonotone Folgerungsrelationen behandeln, zu verstehen, betrachten wir eine axiomatische Darstellung der klassischen Aussagenlogik.[6] Eine solche zerfällt in zwei Blöcke:

A1. logische Axiome, und
A2. logische Regeln.

Die Axiome wollen wir hier nicht als Schemata, sondern als konkrete Formeln betrachten. Zu den Regeln gehört dann die bekannte Einsetzungsregel. Die Blöcke A1 und A2 erzeugen induktiv die Menge der klassisch ableitbaren Formeln (Theoreme) sowie die Relation $\vdash \subseteq \wp(\mathrm{FML}) \times \mathrm{FML}$ der klassischen Ableitbarkeit aus Annahmen (deren Elemente wir Sequenzen nennen).

Nun können wir A1+A2 auf zweierlei Weise material erweitern, nämlich durch

M1. materiale Axiome (einzelne Formeln), und durch
M2. materiale Regeln (aus einzelnen Formeln bestehend).

Wichtig ist nun, daß die Einsetzungsregel (unter A2) nur angewandt wird auf Formeln, die allein aus logischen Axiomen (A1) abgeleitet wurden; anderenfalls greift Satz 2! Wir werden uns zunächst mit dem *vorsichtigen* Gebrauch von M1, so wie in dem Beispiel oben angedeutet, beschäftigen. Statt Annahmen (das Auto ist unmanipuliert) können wir auch Regeln im Hintergrund fixieren; zum Beispiel die Regel: Aus der Annahme, der Schlüssel sei

[6] Ein im wesentlichen gleiches Prinzip liegt [200] zugrunde.

gedreht, schließe, daß der Motor anspringt. Wie wir sehen werden, können sich M1 und M2 verschieden verhalten. Deshalb ist dem vorsichtigen Umgang mit M2 ein eigener Teil dieses Kapitels gewidmet.

Die Menge der Theoreme ist richtig und vollständig hinsichtlich der Menge der Formeln, die unter jeder Interpretation der Atome wahr sind (Tautologien). Genauso stimmt die Relation der Ableitbarkeit aus A1 und A2 überein mit der Relation der Folgerung in Modellen: A folgt aus X genau dann, wenn *jede* Interpretation, die alle Formeln in X wahr macht, auch A wahr macht. Materiale Folgerung entsteht, wenn Interpretationen, die bestimmte, nicht tautologische Formeln falsch machen, gar nicht betrachtet werden, d.h. nicht in den Bereich des Quantors "jede" in der Definition der Folgerung fallen. Diese Modelle unterstellen gewissermaßen die Wahrheit der in den Formeln ausgedrückten materialen Behauptungen. Wie man den vorsichtigen Umgang mit solchen Unterstellungen auf der Modellebene verstehen kann, wird uns zum Schluß beschäftigen.

2. Materiale Annahmen

Poirot untersucht eine merkwürdige Mordserie.[7] Zu einer Reihe von selbstverständlichen Hintergrundannahmen gehörte auch, daß Sir Charles als Verdächtiger eigentlich nicht in Frage kommt und darüber hinaus auch ein Alibi hat (P). Aber nun zieht Miss Milray das Alibi zurück und Poirot muß seine weiteren Schlüsse unter der Annahme $\neg P$ ziehen. Wie sollte Poirot aus $\neg P$ und anderen Prämissen schließen? Wenn er auf seiner Hintergrundannahme P unanfechtbar beharrt, dann schließt er jetzt aus inkonsistenten Annahmen, d.h. (klassisch) auf alles mögliche. Besser: Er sollte erkennen, daß die Frage, was aus $\neg P$ folgt, nur sinnvoll beantwortet werden kann, wenn er die bisherige Hintergrundannahme P jetzt ausklammert. Diese Annahme sollte er deaktivieren und fragen, was aus $\neg P$ vor dem Hintergrund der mit $\neg P$ konsistenten Annahmen folgt.

Allgemein betrachtet, stehen wir vor folgendem Problem. Die Hintergrundannahmen X sind inkonsistent mit einer Formel A, die in einem Schluß als Prämisse verwendet werden soll. Wir fragen jetzt danach, wie wir die Menge $X \cup \{A\}$ so zurückschneiden können, daß aus dem Resultat möglichst viele, sichere und wertvolle Informationen gefolgert werden können. Das Problem ist, daß die drei Parameter *viel*, *wertvoll* und *sicher* in verschiedene Richtungen ziehen. Die inkonsistente Menge $X \cup \{A\}$ ist maximal im Hinblick auf den Parameter viel und minimal im Hinblick auf Sicherheit. Je kleiner die Menge, desto sicherer, aber auch desto wertloser und unergiebiger

[7] In Agatha Christie, *Three Act Tragedy*, dt. *Nikotin*.

ist sie. (Die sicherste Menge ist die leere Menge, aus der genau die Tautologien gefolgert werden können.) Schließlich können Mengen, die in Sicherheit und Größe ununterscheidbar sind, im Hinblick auf ihren Wert weit auseinandergehen. Wir stehen hier vor genau drei theoretischen Optionen.

2.1 Erste Option: Der volle Schnitt.

Poirot spielt auf Nummer Sicher. Er betrachtet alle maximal konsistenten Teilmengen seiner Annahmen und folgert nur aus ihrem gemeinsamen Kern.

DEFINITION 4. (*Folgerung nach dem vollen Schnitt*). Eine Formel A folgt aus einer Formelmenge X vor dem Hintergrund einer Formelmenge K nach dem vollen Schnitt gdw $K', X \vdash A$, für jede Teilmenge $K' \subseteq K$, welche maximal konsistent mit X ist. (Notation: $X \mid\!\sim_K A$ bzw. $A \in \mathrm{C}_K(X)$.)

Die Hintergrundannahmen einer Folgerung werden wir ab jetzt als *Defaults* bezeichnen. Das ist ein treffender und einprägsamer *terminus technicus* aus der englischsprachigen Literatur – es gibt praktisch keine andere –, der auch in anderen Sprachen üblich geworden ist. Nach der Definition sagen wir also auch: A folgt aus Annahmen X und Defaults K nach dem vollen Schnitt.

Erinnerung: Eine Menge X' ist *maximal konsistent* in einer Menge X mit einer Menge Y (maximal Y-konsistent) gdw

1. $X', Y \nvdash \bot$ (konsistent mit Y);
2. $X' \subseteq X$ (in X);
3. $\forall X''$: wenn $X' \subset X'' \subseteq X$, dann $X'', Y \vdash \bot$ (maximal).

Man beachte, daß in der Definition 4 $\vdash$ die deduktive Folgerungsrelation bezeichnet. Wenn Y leer ist (oder nur aus der Konstanten $\top$ besteht), dann reden wir von maximal konsistenten Mengen in X. Wenn X die Menge *aller* Formeln ist, dann reden wir von maximal konsistenten Formelmengen schlechthin. In Definition 4 halten wir die Annahmenmenge X fix und suchen die Defaults in K konsistent mit X zu maximieren.

Wir wollen die Familie aller maximal X-konsistenten Teilmengen einer Menge K mit K_X notieren. Dann können wir die Definition 4 auch so formulieren:

(D4) $$\mathrm{C}_K(X) = \bigcap \{\mathrm{Cn}(K' \cup X) : K' \in K_X\}$$

bzw. so:

$$X \mid\!\sim_K A \text{ :gdw } \forall K' \in K_X : K', X \vdash A.$$

Anmerkung: Macht es einen Unterschied, ob wir, wie in (D4), zuerst unter Konsequenz abschließen und dann schneiden oder, umgekehrt, erst schneiden und dann die Konsequenzen nehmen, wie in:

$$(\mathrm{D}') \qquad \mathrm{Cn}(\bigcap K_A \cup X) \ ?$$

Etwas vereinfacht, fragen wir hier, ob $\mathrm{Cn}(A \cap B) = \mathrm{Cn}(A) \cap \mathrm{Cn}(B)$. Daß die Inklusion von links nach rechts gilt, folgt unmittelbar aus der Monotonie von Cn: Da $A \cap B \subseteq A$, muß $\mathrm{Cn}(A \cap B) \subseteq \mathrm{Cn}(A)$, und da das Gleiche für B gilt, haben wir $\mathrm{Cn}(A \cap B) \subseteq \mathrm{Cn}(A) \cap \mathrm{Cn}(B)$. Die umgekehrte Richtung gilt jedoch nicht, wie ein einfaches Gegenbeispiel zeigt: $P \in \mathrm{Cn}(P \wedge Q) \cap \mathrm{Cn}(P \wedge R)$; aber $\{P \wedge Q\} \cap \{P \wedge R\} = \emptyset$ und sicher nicht $P \in \mathrm{Cn}(\emptyset)$.
Wenn wir Konsequenzen schneiden – wie in (D′) –, dann erhalten wir also mindestens soviel Information, wie wenn wir Schnitte aus Konsequenzen – wie in der Definition (D4) – bilden. Das gerade angegebene Beispiel zeigt, wie wir mit der Anordnung $\mathrm{Cn} \bigcap$ leer ausgehen können, während $\bigcap \mathrm{Cn}$ noch Information hergibt. Von den beiden Optionen ist also letztere die stärkere. In der Theorie der Überzeugungsänderung wird jedoch der ersten, d.h. $\mathrm{Cn} \bigcap$ der Vorzug gegeben, aus Gründen, die wir später erörtern werden. Unter einer prominenten Bedingung – von der die klassische Theorie der Überzeugungsänderung immer ausgeht – koinzidiert die rechte Seite von (D4) mit (D′): dann nämlich, wenn K selbst unter Cn abgeschlossen ist.

Worin besteht nun der wesentliche Unterschied zwischen der materialen Relation $\vdash_K$ und der vorsichtigen Relation $\mid\!\sim_K$ nach dem vollen Schnitt? Erstens, im Gegensatz zu $\vdash_K$, stehen die Defaults K in $\mid\!\sim_K$ für das Schließen nicht vollständig, sondern in der Regel nur teilweise zur Verfügung. Zweitens, und wichtiger, welche Defaults zur Verfügung stehen – ein- bzw. ausgeschaltet sind –, hängt von der jeweiligen Prämissenmenge X ab. Das vereitelt unmittelbar Monotonie: Mit Prämissen X mögen Teile von K zum Schließen verwendet werden, die für eine Obermenge $X' \subseteq X$ ausgeschaltet sein müssen; d.h. nichts garantiert, daß wir aus $X \mid\!\sim_K A$ auf $X' \mid\!\sim_K A$ schließen dürfen.

An einem Beispiel wollen wir uns verdeutlichen wie das Folgern nach dem vollen Schnitt funktioniert.

Beispiel: Es sei

$$K = \{P \to Q, Q \to \neg P, R\}.$$

Was folgt aus P mit Defaults K?

- Unvorsichtig materiale Folgerung Cn_K: Da K inkonsistent mit P ist, haben wir $\mathrm{Cn}_K(P) = \mathrm{Cn}(K \cup \{P\}) = \mathrm{Cn}(\bot) = \mathrm{FML}$.
- Um $\mathrm{C}_K(P)$ zu bestimmen, stellen wir zunächst fest, daß $K_{\{P\}}$ aus diesen zwei Mengen besteht:

$$K_1 = \{R, P \to Q\} \quad \text{und} \quad K_2 = \{R, Q \to \neg P\}.$$

  Nach der Definition 4 ist somit

$$\mathrm{C}_K(P) = \mathrm{Cn}(K_1 \cup \{P\}) \cap \mathrm{Cn}(K_2 \cup \{P\}) \; [\, \supseteq \mathrm{Cn}(P, R) \supset \mathrm{Cn}(P)\,].$$

  $\mathrm{C}_K(P)$ liegt also "in der Mitte" zwischen $\mathrm{Cn}(P)$ und $\mathrm{Cn}_K(P)$.
- Wir haben ferner $R \in \mathrm{C}_K(P)$, jedoch offensichtlich nicht $R \in \mathrm{C}_K(P, \neg R)$. (Um jetzt $K_{\{P,\neg R\}}$ zu bestimmen, müssten wir R aus K_1 und K_2 herausnehmen.) Hier zeigt sich die fehlende Monotonie von C_K. Welchen (konsistenten) Gebrauch wir von den Defaults machen können, hängt eben von den Prämissen ab, aus denen wir schließen wollen.

Das Beispiel illustriert was ganz allgemein für C_K *vis-à-vis* Cn und Cn_K gilt, nämlich

$$\mathrm{Cn} \leq \mathrm{C}_K \leq \mathrm{Cn}_K.$$

(Das wollen wir als Kurzschreibweise vereinbaren für: $(\forall X)\ \mathrm{Cn}(X) \subseteq \mathrm{C}_K(X) \subseteq \mathrm{Cn}_K(K)$.) Die linke Inklusion liegt einfach an der Monotonie von Cn. Denn für für beliebige Mengen K' (und X) ist $\mathrm{Cn}(X) \subseteq \mathrm{Cn}(K' \cup X)$ und daher $\mathrm{Cn}(X) \subseteq \bigcap_{K'} \mathrm{Cn}(K' \cup X)$ $[= \mathrm{C}_K(A)]$. Die rechte Inklusion folgt aus der Beobachtung, daß für jede Menge $K' \subseteq K$ natürlich auch $\mathrm{Cn}(K' \cup X) \subseteq \mathrm{Cn}(K \cup X)$, d.h. $\bigcap_{K' \subseteq K} \mathrm{Cn}(K' \cup X) \subseteq \mathrm{Cn}(K \cup X)$. Daß die Inklusion nicht strikt sein muß, zeigt der Grenzfall, in dem X konsistent mit K ist. Dann gibt es nur eine Teilmenge von K, die *maximal* konsistent mit X ist, nämlich K selbst. Deshalb haben wir in diesem Fall $\mathrm{C}_K(X) = \mathrm{Cn}_K(X)$. Auch $\mathrm{Cn}(X) = \mathrm{C}_K(X)$ ist möglich: Sei $K \subseteq X$ oder $K_X = \emptyset$.

Von den oben erwähnten Eigenschaften weist $\mid\!\sim_K$ (bzw. C_K) die folgenden auf:

Supraklassisch $\qquad \vdash \subseteq \mid\!\sim_K$

Reflexiv $\qquad X, A \mid\!\sim_K A$

Schnitt $\qquad \dfrac{X \mid\!\sim_K \forall Y \qquad X, Y \mid\!\sim_K B}{X \mid\!\sim_K B}$

Disjunktiv $\qquad \dfrac{X, A \mid\!\sim_K C \qquad X, B \mid\!\sim_K C}{X, A \vee B \mid\!\sim_K C}$

Statt Monotonie erfüllt $\mid\!\sim_K$ eine schwächere Form, nämlich

Vorsichtige Monotonie $$\frac{X \mid\!\sim_K A \quad X \mid\!\sim_K B}{X, A \mid\!\sim_K B}.$$

Diese Regel genügt, um aus den übrigen weiter *Rechte Abschwächung* und *Linke Logische Äquivalenz* abzuleiten. (Die Nachweise seien dem Leser zur Übung überlassen.) Die Relation $\mid\!\sim_K$ ist nicht kompakt.

Ein Gegenbeispiel aus [200, p. 33] sei hier kurz skizziert. Es sei X eine Menge von Atomen (endlich oder unendlich). Wir fixieren daraus ein beliebiges Atom P und definieren:

$$K_1 = \{x \wedge P : x \in X\} \quad \text{und} \quad K_2 = \{\neg x \wedge \neg P : x \in X\};$$

$K = K_1 \cup K_2$. Dann ist K_1 die einzige maximal X-konsistente Teilmenge von K und wir haben $K_1, X \vdash P$, d.h. $X \mid\!\sim_K P$. Nun zeigen wir, daß es *keine* (endliche oder unendliche) echte Teilmenge Y von X geben kann mit $Y \mid\!\sim_K P$. Angenommen Y ist eine solche Menge, dann gibt es ein $Q \in X$ so, daß $Q \notin Y$. Wir betrachten die Menge

$$K^* = \{\neg y \wedge \neg P : y \notin Y\}.$$

Dann ist $\neg Q \wedge \neg P \in K^*$ und da K^* offensichtlich konsistent mit Y ist, gilt insbesondere $K^*, Y \nvdash P$. Es läßt sich leicht zeigen, daß K^* eine *maximal* Y-konsistente Teilmenge von K ist. Also $Y \not\mid\!\sim_K P$.

Für alle *unendlichen* Prämissenmengen X gilt jedoch eine schwache Form von Kompaktheit, nämlich

Redundanz $$\frac{X \mid\!\sim_K A}{\exists X' \subset X : X' \mid\!\sim_K A}.$$

Die Definition des Schließens nach dem vollen Schnitt gibt eine Antwort auf zwei wesentliche Fragen:

– Welche Teile der Default-Annahmen sollen im Hinblick auf die Prämissen zum Schließen verwendet werden? Und:
– Falls mehrere Teile in Frage kommen, wie sollen diese Teile aggregiert werden?

Die Antwort der Definition auf die erste Frage ist maximal: Wir betrachten *alle* Teilmengen der Annahmen, die mit den Prämissen maximal konsistent sind. Die Antwort auf die zweite Frage ist gewissermaßen minimal (oder "skeptisch"): Wir betrachten nur den gemeinsamen Kern (den vollen Schnitt) aller dieser Teilmengen. Diese Strategie hat *prima facie* folgendes für sich: Die erste Antwort macht maximalen Gebrauch von der zur Verfügung stehenden Information; die zweite Antwort minimiert das Risiko auf Falsches zu schließen. Das hört sich zunächst gut an, ist aber am Ende doch nicht so gut, wie wir jetzt sehen werden.

Wenn man bestimmte Annahmen bis auf weiteres für richtig hält und deshalb aus ihnen folgert, dann sollte es gleichgültig sein, wie man diese Annahmen formuliert. Jede Formulierung mit dem gleichen Folgerungspotential, d.h. alle logisch äquivalenten Formulierungen der Defaults sollten für die Zwecke des Schließens im Prinzip – wenn vielleicht auch nicht unter praktischen Gesichtspunkten – gleich gut sein. Auch ein anfechtbarer Schluß wäre zurecht kritisierbar, wenn die Konklusion nur in *einer* von vielen logisch äquivalenten, syntaktischen Varianten folgt. Ein solches "Folgern" würden wir in den Bereich fragwürdiger Rhetorik verweisen.

Diese einfache Überlegung spricht dafür, daß das Folgern aus Defaults K die Bedingung erfüllen sollte, daß wenn $K \equiv H$, dann $C_K(X) = C_H(X)$; relational notiert:

Default-Äquivalenz $$\frac{K \equiv H \qquad X \mid\!\sim_K A}{X \mid\!\sim_H A}\,\text{DLE.}$$

Aber das Schließen nach dem vollen Schnitt erfüllt diese Bedingung nicht.

Beispiel: Es sei $K = \{P, \neg Q\}$ und $H = \{P \vee Q, \neg Q\}$. Dann $K_Q = \{\{P\}\}$ und $H_Q = \{\{P \vee Q\}\}$. Daher $Q \mid\!\sim_K P$, jedoch nicht $Q \mid\!\sim_H P$.

Es ist nicht schwierig, die Definition von C_K so zu modifizieren, daß DLE gilt. Dazu müssen wir sicherstellen, daß wir in der Bestimmung derjenigen Defaults, die mit den Annahmen verträglich sind, uns nicht von der syntaktischen Kodierung der Defaults K abhängig machen, sondern auf ihr volles logisches Potential, nämlich auf $\mathrm{Cn}(K)$ zugreifen können. Der Ort in der Definition für die in diesem Sinne richtige Maßnahme ist die Bildung der maximal konsistenten Mengen. Diese bilden wir jetzt nicht aus K, sondern aus dessen logischen Abschluß, $\mathrm{Cn}(K)$:

$$(*) \qquad \mathrm{C}^*_K(A) = \bigcap\{\mathrm{Cn}(K' \cup A) : K' \in \mathrm{Cn}(K)_A\}.$$

Das ist offensichtlich der Spezialfall der ursprünglichen Definition 4, in der wir nun $\mathrm{Cn}(K)$ für die allgemeinere Variable K eingesetzt haben. Etwas anders gesagt, die in $(*)$ definierte Operation C^* ist genau die in (D4) definierte

Operation C, soweit diese auf logisch abgeschlossene Defaultmengen K, d.h. $K = \mathrm{Cn}(K)$, angewandt wird:

$$C^*_K = C_{\mathrm{Cn}(K)}.$$

Übung: Man zeige, daß $\mathrm{C}_{\mathrm{Cn}(K)}$ (wie in Def. 4) DLE erfüllt.

Wenn wir aus Annahmen X unter Defaults K schließen, dann gibt es grundsätzlich immer zwei Möglichkeiten: Die Defaults sind konsistent mit den Annahmen oder sie sind es nicht. Im ersten Fall, können und sollten wir uns aller Defaults bedienen und einfach deduktiv aus $X \cap K$ schließen. Der zweite Fall ist der interessantere: Jetzt sollten wir klug aus den Defaults wählen. Wenn wir jedoch C_K so wie in $(*)$ modifizieren, dann erhalten wir eine uninteressante Folgerungsoperation C^*_K, die im interessanten Fall alle Defaults ignoriert und auf das deduktive Schließen allein aus den Annahmen X zurückfällt. Mit anderen Worten, C^*_K ist in beiden Fällen eigentlich eine deduktive Operation. Das ist der Inhalt des nächsten Satzes.

Satz 5. *Es sei* $K = \mathrm{Cn}(K)$. *Dann* $\mathrm{C}_K(X) = \mathrm{Cn}(K \cup X)$, *falls* X *mit* K *konsistent ist; anderenfalls* $\mathrm{C}_K(X) = \mathrm{Cn}(X)$.

Beweis. Im ersten Fall ist $K_X = \{K\}$ und die Aussage folgt unmittelbar aus (D4).

Für den zweiten Fall, d.h. unter der Annahme

$$K, X \nvdash \bot \tag{1}$$

betrachten wir zunächst die $\supseteq$-Richtung, d.h. $\mathrm{Cn}(X) \subseteq \mathrm{C}_K(X)$. Dies ist nichts anderes als die (richtige) Beobachtung, daß C_K supraklassisch ist.

Wir kommen zur umgekehrten Richtung, $\mathrm{C}_K(X) \subseteq \mathrm{Cn}(X)$. Unter der weiteren Annahme (2) $X \nvdash A$, wollen wir zeigen, daß $X \not\mid\!\sim_K A$. Dazu genügt es, eine Menge H in K_X (den maximal X-konsistenten Teil von K) zu finden so, daß $H, X \nvdash A$.

Aus (1) folgt (Kompaktheit!), daß es eine endliche Menge $X_0 \subseteq X$ gibt so, daß $K, X_0 \vdash \bot$. Es sei B die Konjunktion aller Formeln in X_0. Dann haben wir nach (1), $K, B \vdash \bot$, also $K \vdash \neg B$ und so auch (3) $K \vdash \neg A \vee \neg B$. Das bedeutet nach der Voraussetzung des Satzes, daß (4) $\{\neg A \vee \neg B\}$ eine Teilmenge von $\mathrm{Cn}(K)$ ist.

Aus (2) folgt, daß $X, \neg A \vee \neg B \nvdash \bot$. Denn anderenfalls hätten wir $X \vdash A \wedge B$ und so $X \vdash A$, was (2) widersprechen würde. Das bedeutet, daß (5) $\{\neg A \vee \neg B\}$ konsistent mit X ist.

Aus (4) und (5) folgt nun, daß $\{\neg A \vee \neg B\}$ eine X-konsistente Teilmenge von $\mathrm{Cn}(K)$ ist. Nach Lindenbaums Lemma (oder einem äquivalenten Maximierungsprinzip, wie Zorns Lemma) kann $\{\neg A \vee \neg B\}$ zu einer Teilmenge H von $\mathrm{Cn}(K)$ erweitert werden, die maximal X-konsistent ist; d.h. $H \in \mathrm{Cn}(K)_X$.

Wir wissen, daß $H \vdash \neg A \vee \neg B$ und also auch $H, X \vdash \neg A \vee \neg B$. Wenn nun – *reductio!* – $H, X \vdash A$, dann $H, X \vdash \neg B$. Aber B ist die Konjunktion der Formeln in X. Also wäre H nicht X-konsistent – Widerspruch! So folgt $H, X \nvdash A$, wie gewünscht. ∎

Wir fassen zusammen: Schließen nach dem vollen Schnitt erfüllt nicht die wichtige Äquivalenz-Bedingung DLE. Eine Modifikation würde die Bedingung zwar erfüllen, liefe aber darauf hinaus, daß im interessanten Fall der eigentlich beabsichtigte materiale Schluß die Defaults völlig unberücksichtigt läßt und auf einen deduktiven (formalen) Schluß aus den Prämissen zurückfällt. Verzichten wir auf die Modifikation – und damit auf DLE – müssen wir uns dennoch vor dem – eigentlich wünschenswerten – Fall hüten, daß die Defaults unter logischer Konsequenz abgeschlossen sind, also, im logischen Sinne, eine *Theorie* darstellen. Denn wenn die Default-Theorie nicht völlig verträglich mit den Prämissen ist (der interessante Fall!), dann wird sie für das Schließen aus diesen Prämissen überhaupt nicht berücksichtigt. Wie auch immer, die Konstruktion nach dem vollen Schnitt bringt nicht das gewünschte Ergebnis. Wir müssen daher nach einem besseren Verfahren suchen, die mit den Prämissen konsistenten Teile der Defaults zu aggregieren.

2.2 Andere Strategien für das Schließen aus Defaults.

Die Idee hinter dem vollen Schnitt war eigentlich nicht schlecht. Erstens, ist es nicht sinnvoll aus Prämissen vor dem Hintergrund bestimmter Annahmen zu schließen, ohne darauf zu achten, ob die Annahmen von den Prämissen nicht in Frage gestellt werden. Materiales Schließen sollte vorsichtig vorgehen. Statt beliebige Prämissenmengen mit den immer gleichen Defaults zu kombinieren, sollten wir die zu verwendenden Annahmen an die jeweiligen Prämissen anpassen.

Zweitens, stellen Prämissen und Defaults potentiell wertvolle Information bereit. Solche Information sollten wir nicht ohne guten Grund verwerfen. Vielmehr sollten wir versuchen, die Defaults K so an die Prämissen anzupassen, daß jene das Folgern mit dem unter den Umständen größtmöglichen Informationsgehalt unterstützen. Deshalb sollten wir für das Folgern aus den Prämissen nicht irgendwelche, mit den Prämissen konsistente Teilmengen der Defaults betrachten, sondern nur solche, die diese Eigenschaft in maximaler Weise haben.

Nun gibt es aber möglicherweise viele maximal prämissenkonsistente Teilmengen einer Defaultmenge K. (Für eine Prämissenmenge X haben wir diese Mengenfamilie mit K_X notiert.) Deshalb brauchen wir, drittens, ein nichtwillkürliches Verfahren f, aus K_X eine “gute” Menge $f(K_X)$ zu bilden, aus der wir schließlich folgern wollen.

Es ist klar, daß f nicht die Vereinigung aller Mengen in K_A sein kann. Daß würde zwar Informationsgehalt maximieren, jedoch leider nur in einem trivialen Sinne. Denn für jedes Paar von Elementen in K_A ist deren Vereinigung inkonsistent.

Ein beliebiges Element aus K_A herauszupicken, wäre willkürlich. Denn da K_A keine Äquivalenzklasse unter irgendeiner hier relevanten Äquivalenzrelation ist, kann ein beliebiges Element aus K_A diese Klasse nicht bloß – *pars pro toto* – repräsentieren. Man könnte sich zu der Annahme verhelfen, daß es in jeder Mengenfamilie K_A immer eine optimale Menge gibt, aus der wir schließen sollten. Aber, was immer man unter “optimal” sinnvoll verstehen möchte, diese Annahme ist nicht nur *per se* wenig plausibel, sondern generiert auch – wie wir gleich beobachten werden – mehr Folgerungen, als es uns lieb sein kann. In der Literatur wird eine Konstruktion unter dieser Annahme *Maximalwahl* (*maxi-choice*) genannt.

Wenn kein Element in K_A als bestes ausgezeichnet werden kann, dann liegt das “demokratische” Verfahren nach dem vollen Schnitt nahe: Jede Stimme in K_A soll zählen. Das demokratische Verfahren übersetzt sich hier aber zu einem skeptischen – *zu* skeptisch, wie wir gesehen haben: Im interessanten Fall bleiben nur noch die deduktiven Folgerungen aus den gegeben Prämissen übrig.

In der Mitte zwischen Maximalwahl und vollem Schnitt liegt der sogenannte *Teilschnitt* (*partial meet*). Die Vorstellung, daß unter bestimmten Gesichtspunkten, einige Elemente in K_X anderen vorzuziehen seien, ist sicher nicht abwegig. Die Annahme einer solchen, irgendwie bestimmten Präferenzordnung in K_X garantiert nicht, daß es für beliebige K und X in K_X immer ein einziges maximales Element gibt (d.h. maximal unter der angenommenen Präferenzordnung). Die Annahme garantiert jedoch die Existenz von *mindestens* einem maximalen Element. Daher liegt das folgende, “aristokratische” Verfahren (nur die “Besten” sollen entscheiden) nahe: Sei $s(K_X)$ die Menge der besten Elemente in K_X. Dann soll das Aggregat $f(K_X)$ die Formelmenge $\bigcap s(K_X)$ sein.

Anmerkung zu Auswahlfunktionen und Präferenzordnungen. Im letzten Abschnitt haben wir davon gesprochen, daß wir Elemente in einer Menge aufgrund einer Präferenzordnung auswählen. In dem Ausdruck $s(K_X)$ bildet die *Funktion s* die Menge K_X auf eine Teilmenge von K_X ab; diese

Teilmenge soll die Menge der minimalen Elemente unter der *Relation* ist-besser-als sein. Hier ist ein oft anzutreffendes Verhältnis angedeutet zwischen einer *Auswahlfunktion* s einerseits, und einer *Präferenzrelation* $<$ andererseits. Eine Auswahlfunktion auf einer Menge M (in unserem Fall die Menge der Formelmengen) ist eine Abbildung $s : \wp M \longrightarrow \wp M$ so, daß für jede nichtleere Menge $X \subseteq M$ die Inklusionsbedingung gilt: $s(X) \subseteq X$. Eine Präferenzrelation $< \subseteq M \times M$ ist eine wohlfundierte Relation. "Wohlfundiert" heißt, daß es in jeder nichtleeren Teilmenge X von M minimale Elemente unter $<$ gibt:

$$\forall X \subseteq M : X \neq 0 \Rightarrow \exists x \in X \neg \exists y \in X : y < x.$$

Typischerweise werden Auswahlfunktionen und Präferenzrelationen unter weitere Bedingungen gestellt (die hier jetzt nicht interessieren). Jedenfalls garantiert die Inklusionsbedingung für s, daß es für jede Menge $X \subseteq M$ eine nichtleere Menge $s(X)$ "präferierter" Elemente gibt genauso wie die Wohlfundiertheit von $<$ dafür sorgt, daß es in jeder Menge $X \subseteq M$ "ausgewählte" Elemente gibt. Solange s und $<$ "zueinander passen", können wir zwischen Funktion und Relation frei hin- und hergehen, denn

$$x \in s(X) \text{ gdw } x \text{ ist minimal unter } <, \text{ und } x < y \text{ gdw } x \in s(\{x, y\}).$$

2.3 Zweite Option: Maximalwahl.

Poirot setzt auf die beste Karte. Er betrachtet alle maximal konsistenten Teilmengen seiner Annahmen und sucht die beste aus. Daraus zieht er dann seine Schlüsse.

Definition 6. (*Maximalwahl-Folgerung*) $C_K(X) = \mathrm{Cn}(K' \cup X)$, für eine ausgewählte Formelmenge $K' \in K_X$.

Wir haben oben für die Bedingung DLE argumentiert, derzufolge logisch äquivalente Defaults äquivalente Folgerungsrelationen erzeugen sollten. Wenn Maximalwahl-Folgerungen diese Bedingung erfüllen sollen, dann sollten wir das maximale Element – ähnlich wie bei den vollen Schnitten – aus $\mathrm{Cn}(K)_X$ wählen. D.h. wir sollten die Definition grundsätzlich unter der Bedingung $\mathrm{Cn}(K) = K$ betrachten. Aber dann, so stellt sich heraus, ist die Menge dessen, was sich aus einer beliebigen Prämissenmenge X folgern läßt, immer *vollständig*: Für jede Formel A haben wir entweder $A \in \mathrm{C}_K(X)$ oder $\neg A \in \mathrm{C}_K(X)$! Das trifft auf Poirot'sches Folgern sicher nicht zu. Anderenfalls würde Poirot aufgrund beliebiger Prämissen für jede Frage, die er sich stellen könnte gleich eine definitive Antwort finden. (War Sir Charles der Täter? Aus *jeder* Menge von Prämissen würde eine Antwort auf die Frage sogleich folgen.)

Satz 7. *Es sei* $\mathrm{Cn}(K) = K$. *Dann ist* $\mathrm{C}_K(X) = \mathrm{Cn}(K \cup X)$, *falls* K *mit* X *konsistent ist; anderenfalls ist* $C_K(X)$ *vollständig.*

Beweis. Der erste Teil der Behauptung gilt wie im Falle des Folgerns nach dem vollen Schnitt (siehe Satz 5).

Für den zweiten Teil ("anderenfalls") nehmen wir an, daß (1) $K, X \vdash \bot$. Nach der Definition von $\mathrm{C}_K(X)$ ist nun zu zeigen, daß $K', X \vdash A$ oder $K', X \vdash \neg A$, für ein $K' \in K_X$ und alle Formeln A. Für *reductio* nehmen wir an, daß es eine Formel A gibt so, daß

$$(\dagger) \qquad K', X \nvdash A \text{ und } K', X \nvdash \neg A, \text{ mit } K' \in K_X.$$

Es folgt aus (1) (Kompaktheit), daß es eine Formel B gibt, welche eine endliche Teilmenge von X konjungiert so, daß $K, B \vdash \bot$ und also

$$(2) \qquad K \vdash \neg B, \text{ wobei } X \vdash B.$$

Weiter folgt aus (2), daß $K \vdash A \vee \neg B$ und, nach der Voraussetzung $\mathrm{Cn}(K) = K$, somit (3) $A \vee \neg B \in K$. Auf gleiche Weise erhalten wir auch (4) $\neg A \vee \neg B \in K$.

Angenommen nun, (5) $A \vee \neg B \notin K'$. Dann folgt aus (3) und (†) $K' \in K_X$, daß $K', X, A \vee \neg B \vdash \bot$, d.h. $K', X \vdash \neg A \wedge B$. Und angenommen auch (6) $\neg A \vee \neg B \notin K'$, so folgt auf gleiche Weise aus (4), daß $K', X \vdash A \wedge B$. Also folgt aus (5) und (6), $K', X \vdash \bot$, im Widerspruch zur Annahme (†) $K' \in K_X$. Also ist entweder (5) oder (6) falsch, d.h.

$$(7) \qquad A \vee \neg B \in K' \text{ oder } \neg A \vee \neg B \in K'.$$

Aber (2) $X \vdash B$! Also aus (7), $K', X \vdash A$ oder $K', X \vdash \neg A$ – womit die Annahme (†) wie gewünscht *ad absurdum* geführt ist. ∎

Folgern aufgrund einer Maximalwahl beruht also nicht nur auf der fragwürdigen Annahme, daß es immer genau eine beste prämissenkonsistente Teilmenge der Defaults gibt. Es hat darüberhinaus die Eigenschaft, daß Defaultmengen, die unter logischer Konsequenz abgeschlossen sind, vollständige Folgerungsmengen erzeugen. Aber so funktionieren Poirots Kombinationen nicht. Angenommen Poirot hat seine Prämissen konsolidiert und seine Default-Annahmen entsprechend angepaßt. Wenn alles richtig ist und die Defaults und Prämissen für den Fall ausreichend sind, dann kann er jetzt Sir Charles überführen. Aber selbst Poirot ist nicht so unbescheiden, anzunehmen, daß er nun nur noch seine Defaults unter logischer Konsequenz abschließen muß, um *jede* Frage mit Ja oder Nein beantworten zu können.

2.4 Dritte Option: Teilschnitt.

Poirot betrachtet alle maximal konsistenten Teilmengen seiner Annahmen. Einige davon sind besser als jede andere. Er zieht seine Schlüsse aus dem, worin die besten dieser Teilmengen übereinstimmen.

Definition 8. (*Folgerung nach Teilschnitt*)
$C_K(X) = \bigcap\{\mathrm{Cn}(K' \cup X) : K' \in s_K(K_X)\}$ für eine Menge $s_K(K_X)$ ausgewählter Formelmengen in K_X.

In der Definition ist s_K eine Funktion, die Teilmengen von K auswählt,

$$s_K : K \longrightarrow \wp(\wp\,\mathrm{FML}).$$

Die Auswahlaufgabe ist immer relativ zu einer Defaultmenge K. Es gibt keinen Grund anzunehmen, daß wir aus Information über eine Auswahlfunktion s_K etwas erfahren können über die Wahlen, die eine Funktion s_H $(K \neq H)$ trifft. Das gilt insbesondere auch dann, wenn K und H in der Teilmengenbeziehung zueinander stehen. Auf die Gründe werden wir gleich im Abschnitt über relationale Teilschnitte zurückkommen.

Die Definition ist offenbar eine Verallgemeinerung der beiden vorangegangenen: Voller Schnitt und Maximalwahl sind Grenzfälle der Definition. Im einen Fall ist $s_K(K_X) = K_X$; im anderen Fall steht die Auswahlfunktion s_K unter der Bedingung, daß genau eine Menge in K_X für den (jetzt unnötigen) Schnitt gewählt wird.

Folgerung nach Teilschnitt im weiten Sinne der Definition – d.h. Folgerungen nach vollem Schnitt und nach Maximalwahl einschließend – sind

supraklassisch,
reflexiv,
kumulativ transitiv (unter Schnitt abgeschlossen),
vorsichtig monoton,
disjunktiv, und
linksäquivalent.

Folgerungen nach Teilschnitt im engeren Sinne (also solche, die nicht auf vollem Schnitt oder Maximalwahl beruhen) können auch gefahrlos unter die Bedingung DLE (Default-Äquivalenz) gestellt werden: Im Falle $K = \mathrm{Cn}(K)$ werden nun keine zu kleinen (i.S.v. Satz 5) oder zu großen (i.S.v. Satz 7) Folgerungsmengen erzeugt.

Die Definition 8 geht lediglich von der Annahme aus, daß sich aus den, mit den Annahmen verträglichen, größtmöglichen Defaultmengen irgendwie eine Bestenauswahl treffen läßt. Sie enthält keine Information darüber, wie die Auswahl vor sich gehen soll. Das kann nicht überraschen, denn Poirots

Prüfung der Plausibilität von Defaults in einem gegebenen Fall, kann ihm sicher keine bloß formale Definition abnehmen. Die Auswahl plausibler Defaults setzt immer spezifisches Wissen voraus. Dennoch können wir versuchen, auf die Sache ein wenig mehr Licht zu werfen, indem wir nach formalen Bedingungen fragen, welche jede solche Wahl erfüllen sollte. Dazu bietet es sich an, die Auswahl als Manifestation einer Präferenzrelation zu betrachten.

Relationaler Teilschnitt. Die Konstruktion nach einem Teilschnitt geht offenbar davon aus, daß einige Elemente in K_A "besser" sind als andere. Nun drückt "besser als" eine Relation des Vorziehens aus – eine *Präferenzrelation*, wie wir auch sagen wollen. Ein gutes Modell für eine Präferenz im hier relevanten Sinne ist Quines holistische Vorstellung von Überzeugungsnetzen (*webs of belief*). Überzeugungen sind durch logische und möglicherweise andere inferenzielle Verbindungen miteinander verknüpft. Innerhalb eines Überzeugungsnetzes bilden Überzeugungen daher Pakete, die nur zusammen aufgegeben oder angenommen werden können. Ferner sind solche Überzeugungspakete mehr oder weniger verwundbar durch zuwiderlaufende Evidenz. Einige Pakete von Überzeugungen bilden so etwas wie das Zentrum eines Überzeugungsnetzes. Diese bestehen aus solchen Überzeugungen, deren Bewahrung bei der Konfrontation des Netzes mit konträrer Evidenz der Vorzug gegeben wird; verwundbarere Überzeugungen liegen dagegen an der Peripherie des Überzeugungsnetzwerkes. Das Bild von Zentrum und Peripherie ist auch deshalb gut, weil es nahelegt, daß es sich dabei um einen graduellen Übergang handelt, so daß wir im allgemeinen nur sagen können, daß das eine Überzeugungspaket zentraler (oder nicht peripherer) als das andere, d.h das eine dem anderen vorzuziehen (oder jedenfalls nicht schlechter als das andere) sei.

Das Quine'sche Bild von Überzeugungsnetzen läßt sich gut auf Annahmenmengen übertragen. Im interessanten Fall steht die Annahmenmenge K, aus der wir die Prämissen A ergänzen wollen, in Konflikt mit A. Nicht alle Annahmen in K können für den Zweck des Schließens aus A verwendet werden. Es ist nun naheliegend, sich K mit der Struktur eines Quine'schen Überzeugungsnetzes zu denken. Danach besteht K aus Teilen zusammengehöriger Annahmen. Dabei gibt es Gründe einige Teile anderen vorzuziehen, falls nötig. Das läßt sich sehr einfach durch eine Relation $\leq_K$ zwischen Teilmengen von K modellieren. Wenn wir $\leq$ im Sinne einer Präferenzordnung verstehen wollen, dann sind einige allgemeine Bedingungen für $\leq_K$ recht plausibel. Solche Bedingungen, bis auf eine, sollen uns jetzt nicht interessieren. Was wir hier brauchen, ist lediglich die Bedingung, daß es in beliebigen Teilmengen einer Annahmenmenge immer maximale Elemente gibt:

> Es sei K eine durch $\leq_K \subseteq (\wp(K))^2$ strukturierte (nichtleere) Formelmenge. Dann soll es in jeder (nichtleeren) Teilmenge von $\wp(K)$ mindestens ein $\leq_K$-maximales Element geben.

Diese Bedingung, daß $\leq_K$ in $\wp(K)$ "gedeckelt" ("*stoppered*") sei, erinnert nicht nur zufällig an die Limes-Annahme in Lewis' Semantik für Konditionalsätze; sie erfüllt hier wie dort dieselbe Funktion. Es sei

$$\max K_A = \{K' \in K_A : \forall X \in K_A,\ \text{wenn } K' \leq_K X,\ \text{dann } X = K'\}.$$

D.h. $\max K_A$ ist die Menge der besten maximal A-konsistenten Teilmengen von K. Nun ist es naheliegend danach zu fragen, was aus A zusammen mit dem gemeinsamen Kern der besten, der mit A verträglichen Annahmenmengen, die K zur Verfügung stellt, folgt. Das ist der intuitive Gehalt der folgenden Definition:

$$(*) \qquad \mathrm{C}_K(A) = \bigcap\{\mathrm{Cn}(K' \cup A) : K' \in \max K_A\}.$$

Ersichtlich haben wir hier wieder einen Fall der Definition 8 des Folgerns nach Teilschnitt vor uns. Die Aufgabe der Auswahlfunktion s_K übernimmt nun die Funktion max, welche aufgrund der Ordnung $\leq_K$ definiert wurde.

Es ist klar, daß wenn wir uns eine Defaultmenge K wie ein Quine'sches Überzeugungsnetz denken, die Präferenz für Teilmengen von K immer nur für K gelten kann. Ein Überzeugungsnetz K erhält seine Identität auch durch die Angabe einer Präferenzrelation $\leq$ in $(\wp(K))^2$, d.h. solche Netze sind wesentlich Paare $(K, \leq)$. So sind $(K, \leq)$ und $(K, \leq')$ (mit $\leq \neq \leq'$) verschiedene Überzeugungsnetze bzw. strukturierte Defaultmengen. Gegeben eine Annahmenmenge X sind die Folgerungen vor dem Hintergrund $(K, \leq)$ typischerweise andere als die vor dem Hintergrund $(K, \leq')$.

Wenn nun aber K in K' enthalten ist, sollte dann nicht auch $\leq$ eine Teilrelation von $\leq'$ sein? Daß eine solche Bedingung zu nichts Gutem führt, kann man sich an einfachen Beispielen, wie dem folgenden, deutlich machen.

> *Beispiel*: In einer bestimmten Menge von Hintergrundannahmen K mag der Zeugenaussage P ein besonderes Gewicht zukommen. In unserer Modellierung würde das dadurch zum Ausdruck kommen, daß alle unter $\leq_K$ maximalen Teilmengen von K die Annahme P enthalten, d.h. P ist immer zu berücksichtigen. Aber was geschieht, wenn wir uns K um eine Aussage Q erweitert denken, welche die Aussage P stark erschüttert? In

diesem Fall würden wir nicht nur Q über P stellen, sondern viele andere Aussagen könnten ebenfalls in neuem Lichte erscheinen, d.h. einen neuen Stellenwert bekommen. Die Erweiterung von K um Q mischt die Karten neu: Die Präferenz wird nicht erweitert, sondern neu geordnet.

Das Beispiel weist auf eine weitere Manifestation von Nichtmonotonie beim vorsichtigen Schließen aus Hintergrundannahmen. Die Schlüsse, die sich vor dem Hintergrund einer Defaultmenge K ziehen lassen, bleiben nicht generell erhalten, wenn wir K erweitern. In der relationalen Modellierung des Schließens nach Teilschnitt zeigt sich das darin, daß wir aus $K \subseteq H$ nicht auf $\leq_K \subseteq \leq_H$ schließen dürfen. In der Modellierung mit Hilfe von Auswahlfunktionen ist entsprechend der Schluß auf $s_K(M) \subseteq s_H(M)$ (M sei irgendeine Familie von Formelmengen) ungültig.

2.5 Auswahl nach Informationsgehalt.

Teilschnitte scheinen die goldene Mitte zwischen vollen Schnitten und Maximalwahlen zu treffen. Sie werfen aber ein Problem auf für eine kohärente Beschreibung der Theorie anfechtbaren Schließens. (Dieses Problem stellt sich in ganz ähnlicher Weise in der Theorie der Überzeugungsänderung.) Das Ziel solchen Schließens ist es, auf vorsichtige Weise möglichst viel Information aus Prämissen aufgrund *prima facie* plausibler Hintergrundannahmen (Defaults) zu gewinnen. Je mehr Defaults wir heranziehen, umso mehr können wir aus den Prämissen schließen. Vorsicht ist nur im Hinblick auf die Konsistenz der Annahmen, daß heißt der Prämissen zusammen mit den Defaults, geboten. Wir streben also nach einem Optimum: Die Information, die in den Defaults enthalten ist, soll unter der Bedingung der Konsistenz mit den Prämissen optimal genutzt werden. Anders formuliert: In der Anpassung der Defaults an die Prämissen soll Informationverlust minimiert werden.

Vor diesem Hintergrund ist der erste Schritt aller Teilschnittkonstruktionen völlig natürlich: Wir identifizieren zunächst alle Teilmengen der Defaultmenge K, welche die Optimum-Bedingung (gegeben Prämissen X) erfüllen. In dem bisher besprochenen Ansatz ist das – naheliegenderweise – die Mengenfamilie K_X. Alle und nur die Mengen in K_X minimieren den durch X erzwungenen Informationsverlust. Wenn K_X mehr als eine Menge enthält, müssen wir eine nicht-willkürliche Entscheidung treffen. Vielleicht kommen nach weiteren Überlegungen nur noch einige der Mengen mit minimalem Informationsverlust in die engere Wahl. Daß solche Überlegungen immer zu eindeutigen Ergebnissen führen, ist jedoch eine zu starke Annahme. Wenn K_1 und K_2 die ausgewählten Mengen in K_X sind und es keine weiteren Möglichkeiten mehr gibt, zwischen ihnen nicht-willkürlich zu entscheiden, dann sollten wir beide gleich behandeln und dafür gibt es nur eine Umsetzung: Wir betrachten jetzt den Schnitt $K_1 \cap K_2$. Das Gleichbehandlungsprinzip ist aber nur dann vernünftig, wenn der Schnitt zweier gleich guter Kandidaten nicht schlechter ist als diese ist; anderenfalls wäre es besser eine Münze zu werfen. Nun ist es jedoch so, daß wenn K_1 und K_2 in K_X sind, dann ist $K_1 \cap K_2$ typischerweise nicht in K_X. Hatten wir nicht gerade festgestellt, daß nur die Mengen in K_X den Informationsverlust minimieren und also in diesem Sinne optimal sind? Wenn wir nun $K_1 \cap K_2$ als Default-Basis für unsere Schlüsse aus X wählen, dann geben wir mehr Information auf, als es nötig wäre. Diese Default-Basis ist nicht optimal – jedenfalls nicht in dem Sinne von "optimal", den wir bisher zugrundegelegt haben. Das Gleichbehandlungsprinzip ist hier also nicht vernünftig anwendbar.

Es gibt grundsätzlich zwei Wege, dieses Problem zu lösen. Am einfachsten ist es, zu sagen, daß K_X eine Vorauswahl darstellt und daß wir nach dem optimalen Schnitt von Teilmengen aus K_X suchen. Danach sind Elemente in K_X *per se* gar nicht optimal im hier relevanten Sinne. "Optimal" hat nicht die einfache Bedeutung, eine maximal X-konsistente Teilmenge von K zu sein; die Bedeutung wird vielmehr durch die Definition eines Teilschnitts festgelegt. Damit umgehen wir zwar das genannte Problem, aber die Rede vom minimierten Informationsverlust kann nun die Teilschnittkonstruktion nicht mehr unabhängig begründen. Das Problem wird in der Literatur zwar selten behandelt, aber wir dürfen unterstellen, daß dieser Weg die Standardlösung darstellt.

Der andere Weg besteht darin, die Kandidatenmenge K_X anders zu bestimmen. Dazu beginnen wir damit, Aussagenmengen einen Informationswert zuzumessen. Gegeben der Informationswert von K, sollen die Mengen in K_X diesen Wert maximal approximieren unter der Bedingung, mit X konsistent zu sein. Plausiblerweise werden alle maximal X-konsistenten Teilmengen von K zu K_X gehören, aber möglicherweise noch weitere. So kann optimaler Informationswert, im hier benötigten Sinne, eine Eigenschaft sein, die unter Schnitten abgeschlossen ist. Wenn das so ist, dann läßt sich das Gleichbehandlungsprinzip anwenden: Der Schnitt zweier gleich optimaler Mengen ist ebenfalls optimal. Diese Idee hat Levi [174, 176] in einer Reihe von Varianten bzw. Verbesserungen im Kontext einer Theorie der Überzeugungsänderungen entwickelt. Eine sehr eingehende Diskussion sowie die wichtigsten Resultate finden sich in Arbeiten von Rott und Pagnucco [251, 253]; vgl. auch [200].

3. Drei Beispieltheorien

Wir betrachten hier kurz drei Ansätze zu nichtmonotonem Schließen. Die ersten zwei sind in der Frühphase der Theorienbildung jeweils sehr einflußreich gewesen und finden nun ihren natürlichen Ort unter der Perspektive, die wir oben entwickelt haben. Die dritte Theorie entwickelt die zweite weiter und deutet Anreicherungen der Modelle an so, wie sie typisch sind für den Fortgang der Theoriebildung.

3.1 Die Annahme der geschlossenen Welt. Der erste Ansatz, den wir betrachten wollen, umgeht das Auswahlproblem, indem er immer nur eine Menge garantiert, die maximal konsistent mit den jeweiligen Prämissen verträglich ist. Das ist das Schließen unter der sogenannten Annahme einer geschlossenen Welt (*closed world assumption*, CWA). Diese Art des Schließens wurde in den 70er Jahren von Reiter untersucht und ist damit eines der frühesten und auch einfachsten nicht-probabilistischen Beispiele für nichtmonotones Schließen.

Es mag sein, daß Poirot keinen Hinweis darauf findet, daß sich Sir Charles am Tatort aufgehalten hat. Daraus zu schließen, daß Sir Charles *nicht* am Tatort war, ist hochriskant und schneidet Ermittlungsmöglichkeiten ab, die vorerst im Spiel bleiben sollten. In anderen Fällen jedoch, ist ein Schluß nach dem Muster

Gehe von $\neg A$ aus, wenn es keinen Hinweis auf A gibt!

durchaus empfehlenswert. Zum Beispiel, wenn Poirot auf der Speisekarte kein Fischgericht findet, dann sollte er davon ausgehen, daß Fischgerichte nicht angeboten werden. Oder wenn Poirot feststellt, daß der Fahrplan keinen Zug nach Dover zu einer bestimmten Uhrzeit verzeichnet, dann sollte er daraus schließen, daß zu dieser Zeit kein Zug nach Dover abfährt. Speisekarten und Zugfahrpläne stellen gewissermaßen geschlossene Welten dar. Es ist normalerweise nicht anzunehmen, daß die Information, die in diesen Quellen enthalten ist, aus anderen Quellen vervollständigt wird. Dagegen ist die Evidenzlage in einem Mordfall bis zur endgültigen Klärung des Falls weiter offen. Daß sich ein Mordfall unerwartet entwickelt, damit sollte man rechnen. Daß Züge fahren, die im Fahrplan nicht verzeichnet sind, damit sollte man nicht rechnen. Es ist unmittelbar klar, daß Schließen nach diesem Muster nicht monoton sein kann.

Die Annahme einer geschlossenen Welt ist ein wichtiges Konstruktionsprinzip für Datenbanken. So wäre es ineffektiv für jedes Buch einer Bibliothek, das nicht ausgeliehen ist, in die Datenbank des Ausleihesystems einzutragen, daß es nicht ausgeliehen ist. Stattdessen sollte standardmäßig jedes Buch im Bestand als nicht ausgeliehen gelten, solange das Buch nicht explizit als “ausgeliehen” eingetragen ist. Zwar bedeuten “Es liegt kein Ausleiheeintrag vor” und “Das Buch ist nicht ausgeliehen” sicher nicht dasselbe: Bücher können unverbucht ausgeliehen bzw. wieder eingestellt sein. Aber die Welt einer Bibliothek ist geschlossen genug, um den Schluß von der Abwesenheit eines Eintrags auf das Vorhandensein des Buches gut genug sein zu lassen.

Folgern unter der Annahme einer geschlossenen Welt fordert dazu auf, anzunehmen daß nichts der Fall ist, außer dem, von dem wir explizit annehmen, daß es der Fall ist. Diese Idee läßt sich auf verschiedene Weisen in eine formale Theorie bringen. Hintergrund ist immer eine Menge K negativer Literale (siehe die Erläuterung unten). In der Abwesenheit jeglicher Prämissen, sind alle Elemente in K aktiviert. (Vgl.: In der Abwesenheit von Ausleihvermerken, gelten alle Bücher als nicht ausgeliehen.) Sobald Prämissen X hinzutreten, werden Elemente in K, die mit diesen Prämissen nicht verträglich sind, ausgeschaltet. D.h., wenn $X \vdash P$, dann steht $\neg P$ als Annahme nicht zur Verfügung; anderenfalls bleibt $\neg P$ als Annahme eingeschaltet. So kommen wir zur Definition der Menge $cwa_K(X)$, der Menge der Prämissen X, die nach der CWA erweitert wird (aus K):

$$\text{(cwa)} \qquad cwa_K(X) = \{\neg P \in K : X \nvdash P\} \cup X.$$

Eine Formel A folgt aus X unter der CWA relativ zu K ($A \in \mathrm{C}_K(X)$) genau dann, wenn

$$cwa_K(X) \vdash A; \text{ d.h. } \mathrm{C}_K(X) = \mathrm{Cn}(cwa_K(X)).$$

Das, so stellt sich schnell heraus, ist etwas zu einfach. Es sei $X = \{P \vee Q\}$ und $K = \{\neg P, \neg Q\}$. Da weder $P \vee Q \vdash P$ noch $P \vee Q \vdash Q$, bleiben sowohl $\neg P$ als auch $\neg Q$ aktiv; d.h. $cwa_K(X) = K \cup X$. Aber dann ist $\mathrm{C}_K(X)$ inkonsistent.

Hier beginnen die Variationen des Folgerns unter der CWA. Die Schlüsselfrage ist: Wie können wir garantieren, daß die Anreicherung der Prämissenmenge X nach der CWA konsistent ist, vorausgesetzt, daß X selbst konsistent ist? Es gibt hier zwei Parameter, an denen wir drehen können: die Anreicherungsfunktion cwa_K und die zulässige Prämissenmenge X (der Definitionsbereich von cwa_K). Untersuchungen zur CWA im weiteren Sinne fragen im wesentlichen nach dem Wechselspiel dieser beiden Parameter. Wir wollen hier nur die Frage möglicher Bedingungen für zulässige Prämissenmengen betrachten.

Das oben genannte, für (cwa) unangenehme Beispiel könnte durch die Forderung ausgeschlossen werden, daß Prämissenmengen keine Disjunktionen enthalten dürfen, ohne mindestens eines der Disjunkte zu enthalten.

Aber obwohl dies eine sehr starke Bedingung ist, bietet sie keine Gewähr dafür, daß sie zugleich allgemein genug ist. Die folgende Bedingung, so stellt sich heraus, garantiert die gewünschte Konsistenz:

(H) Die Prämissenmenge X bestehe ausschließlich aus Formeln, die jeweils äquivalent zu einer Horn-Formel sind.

Eine Formelmenge, welche die Bedingung (H) erfüllt, wollen wir eine *Horn-Menge* nennen.

Erläuterung. Ein *Literal* ist ein Atom P (positives Literal) oder die Negation $\neg P$ eines Atoms (negatives Literal). Eine *Horn-Formel* ist eine Disjunktion von Literalen von denen höchstens eines positiv ist. Horn-Formeln sind also von der Form

$$\neg P_1 \vee ... \vee \neg P_n \vee Q,$$

was äquivalent ist zu

$$P_1 \wedge ... \wedge P_n \to Q.$$

Da wir auch die Konstanten $\perp$ und $\top$ als Atome im Aufbau einer Horn-Formel zulassen wollen, schließt die Definition zwei Grenzfälle ein: 1. Jedes $P_i = \perp$ $(1 \leq i \leq n)$; 2. $Q = \top$. Im ersten Fall stellt sich jedes Atom, im zweiten Fall jede Disjunktion negativer Literale und letztlich auch jedes negative Literal als Grenzfall einer Horn-Formel dar. (Eine Horn-Formel, die keinem dieser Grenzfälle zuzurechnen ist, heißt *strikt.*) Das oben angeführte Beispiel $A = \{P \vee Q\}$ ist nach Bedingung (H) ausgeschlossen, da die Formel in A mehr als ein positives Disjunkt enthält.

Eine wichtige Eigenschaft von Horn-Mengen wollen wir im nächsten Lemma festhalten.

Lemma 9. *Es sei X eine Horn-Menge und P, Q seien Atome. Dann gilt die folgende Disjunktionseigenschaft:*

(Disj) *Wenn $X \vdash P \vee Q$, dann $X \vdash P$ oder $X \vdash Q$.*

Beweis. Wir nehmen an, daß $X \nvdash P$ und $X \nvdash Q$ und zeigen, daß $X \nvdash P \vee Q$. Unter den Annahmen gibt es Modelle $\models_1$ und $\models_2$, die jeweils auf Bewertungen I_1 und I_2 der Atome basieren, so daß

(∗) $$\models_1 X \text{ und } \models_2 X,$$

während $I_1(P) = I_2(Q) = 0$. Wir definieren nun eine Bewertung u für alle Atome x im Definitionsbereich von I_1 oder I_2 so:

$$u(x) = 1 \text{ gdw } I_1(x) = I_2(x) = 1.$$

Es sei $\models_u$ die Erweiterung von u zu einem Modell aller Formeln. (Wenn X inkonsistent ist, dann gilt der Satz trivialerweise; anderenfalls gibt es genaue eine solche Erweiterung.) Wir zeigen, daß $\models_u X$ und $\not\models_u P \vee Q$.

1. $\not\models_u P \vee Q$. Denn $I_1(P) = 0$ und also $u(P) = 0$. Ebenso $I_2(Q) = 0$ und also $u(Q) = 0$.

2. $\models_u X$. Wir nehmen an – *reductio*! –, es gäbe eine Formel $A \in X$ mit $\not\models_u A$. Da A Horn ist, ist

$$A = \neg R_1 \vee \cdots \vee \neg R_n \vee R,$$

wobei $R_1, \ldots, R_n, R$ Atome sind. Drei Fälle sind zu betrachten:

(a) $R_i = \bot$ $(\forall i : 1 \leq i \leq n)$, d.h. $A \equiv R$. Dann ist $u(R) = 0$, und also entweder $I_1(R) = 0$ oder $I_2(R) = 0$ – im Widerspruch zu $(*)$.

(b) $R = \top$, d.h. $A = \neg R_1 \vee \cdots \vee \neg R_n$. Dann ist $\not\models_u \neg R_1 \vee \cdots \vee \neg R_n$ und also $u(R_i) = 1$ $(\forall i : 1 \leq i \leq n)$. Aber dann ist $I_1(R_i) = 1$ und also $\not\models_1 A$, daher $\not\models_1 X$ – im Widerspruch zu $(*)$.

(c) Keiner der beiden gerade betrachteten Fälle, d.h. A ist im strikten Sinne Horn. Dann haben wir für alle i, $u(R_i) = 1$ und $u(R) = 0$. Daraus folgt $(\forall i)$, $I_1(R_i) = 1 = I_2(R_i)$ und entweder $I_1(R) = 0$ oder $I_2(R) = 0$. Das bedeutet, daß entweder $\not\vdash_1 X$ oder $\not\vdash_2 X$ – wiederum im Widerspruch zu $(*)$. ∎

Die Prämissenmenge $X = \{P \vee Q\}$ in dem oben angegebenen Beispiel gegen die unbeschränkte Anwendung von (cwa) hat die Disjunktionseigenschaft offenbar nicht. Es ist genau diese Eigenschaft von Horn-Mengen, die es erlaubt, den folgenden Satz zu beweisen.

Satz 10. *Wenn X eine konsistente Horn-Menge ist, dann ist $cwa_K(X)$ konsistent (für beliebige Mengen K negativer Literale).*

Beweis. Wir nehmen an, daß $cwa_K(X)$ inkonsistent und X konsistent sei, und zeigen, daß X dann keine Horn-Menge sein kann.

Angenommen also, $cwa_K(X) \vdash \bot$, d.h. (1) $\{\neg P \in K : X \not\vdash P\}, X \vdash \bot$. Unter der weiteren Annahme $X \not\vdash \bot$ folgt aus (1), daß es $\neg P_1, ..., \neg P_n$ in K gibt, so daß (2) $\neg P_1, ..., \neg P_n, X \vdash \bot$, wobei (3) $X \not\vdash P_i$ (für alle p_i $(1 \leq i \leq n)$). Aus (2) folgt (4) $X \vdash P_1 \vee \cdots \vee P_n$. Wenn nun X eine Horn-Menge ist, dann folgt aus (4) nach (Disj) für mindestens ein P_i, daß $X \vdash P_i$ – im Widerspruch zu (3). Also kann X unter den Annahmen keine Horn-Menge sein. ∎

Wie hängt nun das Folgern nach der CWA zusammen mit dem Folgern nach Schnitten? Der Zusammenhang wird sofort sichtbar, wenn wir die Definition der Folgerns nach der CWA voll ausschreiben,

(C) $$\mathrm{C}_K(X) = \mathrm{Cn}(\{\neg P \in K : X \not\vdash P\} \cup X),$$

und mit Folgerung nach Maximalwahl vergleichen:

(M) $\mathrm{C}_K(X) = \mathrm{Cn}(K' \cup X)$, mit K' maximal in K_X.

Die Form der Definitionen ist gleich. In (C) tritt an die Stelle der aus K_X ausgewählten Menge K' die Menge $\{\neg P : X \not\vdash P\}$ aller mit X konsistenten Literale aus K. Diese Menge ist deduktiv eindeutig bestimmt, weshalb es keiner eigenen Auswahlfunktion bedarf. Das Folgern nach der CWA ist also eine besonders einfache Art des Maximalwahlfolgerns.

Ein Kollaps, wie in den Sätzen 5 und 7 beobachtet, findet beim Folgern unter der CWA gleichwohl nicht statt. Wir erinnern uns: Das unerwünschte Resultat – die Vollständigkeit der Folgerungsmenge – stellt sich ein, wenn wir den Fall betrachten, daß es sich bei den Defaults K um eine Theorie handelt, d.h. $K = \mathrm{Cn}(K)$. Dieser Fall ist eigentlich gewünscht, da nur so das Folgern unabhängig ist von der syntaktischen Form der Defaults. Aber für die CWA ist es wesentlich, daß sie die expliziten Prämissen aus einem Vorrat negativer Literale erweitert. Das bedeutet, daß die Hintergrundannahmen K mit gutem Grund einer syntaktischen Beschränkung gehorchen müssen, die beim Abschluß unter logischer Folgerung verloren gehen würde. Die Bedingung, unter welcher der Kollaps herbeigeführt wird, nämlich der logische Abschluß der Hintergrundannahmen K, kann also für die unter der CWA definierte Folgerungsoperation C_K grundsätzlich nicht erfüllt sein.

3.2 Poole-Folgerung. Aus dem Umstand, daß das Opfer starke, rauhe Hände hat, schließt Poirot gern, daß es sich jedenfalls nicht um einen Angehörigen der Oberschicht handeln kann. Aber der Schluß von den feinen Händen auf die feine Gesellschaft ist nur dann gut, wenn das Opfer nicht passionierter Hobbyholzhacker ist (wie es Wilhelm II. war). Solange wir nicht wissen, daß das Opfer dieser, in der Oberschicht eher seltenen Passion nachging, solange kann das angenommene Konditional in Kraft bleiben. Sobald unser Wissen aber inkonsistent mit der Vorbedingung des Konditionals wird, sollten wir das Konditional nicht zum Schließen verwenden. Neben den Hintergrundannahmen, aus denen wir unsere Prämissen ergänzen, scheint es im Hintergrund also auch Bedingungen zu geben, die darüber wachen, daß wir die Hintergrundannahmen gefahrlos verwenden können.

Besonders riskant in diesem Sinne scheinen konditionale Annahmen zu sein. So stehen in Gebrauchsanweisungen für Kameras Sätze wie: Wenn auf den Auslöser gedrückt wird, dann öffnet sich der Verschluß ($a \to v$). Solch ein Konditional vermittelt wichtiges Hintergrundwissen für die Bedienung einer Kamera. Aber was wollen wir dann aus der Prämisse, daß der Verschluß sich nicht öffnet ($\neg v$), schließen? Da $\neg v$ viele Ursachen haben kann, wäre es sehr riskant, zu schließen, daß der Auslöser nicht gedrückt wurde ($\neg a$). Aber $a \to v$ und $\neg v \to \neg a$ sind logisch äquivalente Aussagen. Modus Ponens nach a aus dem einen Konditional sollte daher genauso gut sein wie Modus Ponens nach $\neg v$ aus dem anderen.

Es gibt eine einfache Möglichkeit, zwischen $A \to B$ und $\neg B \to \neg A$ zu differenzieren. Wenn wir nach A aus $A \to B$ schließen wollen, nicht aber nach $\neg B$ aus $\neg B \to \neg A$, dann können wir diese Schlußmöglichkeit "deaktivieren", indem wir festlegen, daß $\neg(A \to B)$, falls $\neg B$. Poole [219] enthält einen Vorschlag, wie man das "Deaktivieren" von Defaults ganz allgemein verstehen kann. Danach ist das nur ein spezieller Fall eines viel allgemeineren Phänomens: des Schließens unter *kontrollierter* Hinzuziehung von Defaults. Beim Poole-Folgern treten zu den Defaults K noch Bedingungen B hinzu, die darüber wachen, daß die Defaults auch angewandt werden können. Solche, die Defaults kontrollierenden Bedingungen sind ein Leitmotiv beim Schließen mit Default-Regeln nach Reiter, ein Thema, das uns im nächsten Abschnitt beschäftigen wird. In die Definition des Poole-Folgerns ist dieses Leitmotiv in besonders einfacher Weise eingeflochten. Sehen wir uns zunächst diese Definition an:

Definition 11. Eine Formel A ist eine *Poole-Folgerung* aus Prämissen X unter Defaults K mit Bedingungen *Bed* ($X \mid\!\sim_{KBed} A$) gdw $K', X \vdash x$, für alle $K' \in K_{X \cup Bed}$.

Als Operation notiert, ist Poole-Folgerung demnach so definiert:

$$\text{(D11)} \qquad \mathrm{C}_{KBed}(X) = \bigcap\{\mathrm{Cn}(K' \cup X) : K' \in K_{X \cup Bed}\}.$$

Es fällt sogleich die Ähnlichkeit zur Konstruktion (D4) nach dem vollen Schnitt auf:

$$\text{(D4)} \qquad \mathrm{C}_K(X) = \bigcap\{\mathrm{Cn}(K' \cup X) : K' \in K_X\}.$$

Tatsächlich koinzidiert im Falle $Bed \subseteq X$ die in (D11) definierte Operation C_{KBed} mit der Folgerungsoperation nach dem vollen Schnitt; d.h. wenn C' die Operation nach dem vollen Schnitt ist, dann gilt:

$$\text{Wenn } Bed \subseteq X, \text{ dann } \mathrm{C}_{KBed}(X) = \mathrm{C}'_K(X);$$

und insbesondere ist das Folgern nach dem vollen Schnitt offensichtlich nichts anderes als "bedingungsloses" Poole-Folgern, d.h. $\mathrm{C}_{K\emptyset} = \mathrm{C}'_K$.

Die Rolle der Bedingungen *Bed* wird anhand des oben besprochenen Beispiels deutlich. Es sei $X = \{\neg v\}$, $K = \{a \to v\}$ und $Bed = \{\neg v \to \neg(a \to$

$v)\} = \{a \vee v\}$. (*Bed* deaktiviert also den Default in K im Falle $\neg v$, der durch die Prämisse X gegeben ist.) Dann ist $K_{X \cup Bed} = \{\emptyset\}$ und folglich, d.h. nach (D11), ist $C_{KBed} = \mathrm{Cn}(X)$. In dem Beispiel wissen wir also trotz unseres Defaults über Kameraverschlüsse nicht mehr, als daß der Verschluß sich nicht öffnet. Man beachte, daß die Bedingung in *Bed* hierbei nur eine negative Rolle spielt. Sie schränkt den Zugriff auf den Default ein und generiert nicht selbst neue Information aus den Prämissen.

Wenn wir dagegen aus $\neg v$ unter Verwendung von K nach dem vollen Schnitt (ohne Bedingungen) schließen würden, sähe die Sache anders aus. Dann ist $K_X = \{\{a \rightarrow v\}\}$ und also ist $\mathrm{C}_K(X) = \mathrm{Cn}(\{\neg v, a \rightarrow v\})$. In diesem Fall könnten wir allein aufgrund der Prämisse, daß der Verschluß sich nicht öffnet, darauf schließen, daß der Auslöser nicht gedrückt wurde. Ein deutlich riskanterer Schluß als z.B. der Schluß vom gedrückten Auslöser auf den geöffneten Verschluß.

Schließlich wollen wir auch den Fall betrachten, daß wir die Bedingung in *Bed* in die Defaults K aufnehmen und dann nach dem vollen Schnitt schließen. In diesem Fall gehen wir von $K = \{a \rightarrow v, a \vee v\}$ aus und haben dann $K_X = \{\{a \rightarrow v\}, \{a \vee v\}\}$. Da der Schnitt der Mengen in K_X leer ist, ist $C_K(X) = \mathrm{Cn}(X)$ und generiert damit dieselben Folgerungen wie C_{KBed}.

Natürlich ist es nicht allgemein der Fall (wie im letzten Beispiel), daß $\mathrm{C}_{KBed}(X)$ dieselbe Formelmenge ist wie $C_{K \cup Bed}(X)$ (nach dem vollen Schnitt). Während in $\mathrm{C}_{K \cup Bed}$ die Menge *Bed* die Annahmen in K anreichert, d.h. potentiell zusätzliche Prämissen für das Schließen zur Verfügung stellt, ist das in C_{KBed} nicht der Fall. Hier hat *Bed* einen ausschließlich negativen Effekt: Aus *Bed* werden keine zusätzlichen Prämissen entnommen; vielmehr schränkt *Bed*, neben X, die Menge verwendbarer Defaults in K ein. Die Bedingungen generieren nicht potentiell mehr, sondern weniger Folgerungen.

Wir haben oben gesagt, daß der Schluß vom gedrückten Auslöser (a) auf den geöffneten Verschluß (v) deutlich weniger riskant ist als der kontraponierte Schluß von $\neg v$ auf $\neg a$. Gefragt, diese Risikoeinschätzung zu begründen, würden wir typischerweise kontrafaktische Konditionale anführen.

> $a \sqsupset v$: Wenn der Auslöser gedrückt würde, dann würde sich der Verschluß öffnen.
>
> $\neg v \sqsupset \neg a$: Wenn sich der Verschluß nicht öffnen würde, dann wäre der Auslöser nicht gedrückt worden.

Während $a \sqsupset v$ wahr ist, ist $\neg v \sqsupset \neg a$ falsch. Das erklärt, warum $\neg v \sqsupset \neg a$ riskanter als $a \sqsupset v$ ist – und damit konkurrierende Erklärungen sind nicht in Sicht. Wenn wir also zwischen den Annahmen "wenn a, dann v" und "wenn nicht v, dann nicht a" unterscheiden wollen, dann ist es naheliegend, diese Annahmen nicht als materiale, sondern als kontrafaktische Konditionale zu verstehen. Stellt Poole-Folgerung also so etwas wie eine logische Analyse kontrafaktischer Konditionale dar, die zur üblichen Theorie (nach Lewis und anderen) in Konkurrenz steht? (Wenn das so sein sollte, würde sich schnell herausstellen, daß die Poole'sche Theorie in dieser Konkurrenz unterliegen muß.)

Das hier verwendete motivierende Beispiel (das im wesentlichen auch in Pooles Aufsatz [219] bemüht wird) mag diese Vermutung nahelegen. Die Definition der Poole-Folgerung bietet jedoch ein sehr allgemeines Mittel an, zwischen Formeln, die im klassischen Sinne äquivalent sind, einen Keil zu treiben: im obigen Falle zwischen eine materiale Implikation und deren Kontraposition. Das bedeutet, daß durch die Angabe geeigneter Bedingungen zwischen wahrheitsfunktionalen Formeln feiner unterschieden werden kann, als dies in der klassischen Aussagenlogik möglich ist. Man mag das für nicht mehr als die begrenzte Simulation einer besseren, prinzipielleren Analyse halten. Der Vorzug der Poole'schen Definition liegt gleichwohl darin, daß sie etwas leistet, was sonst nur bei Aufgabe der wahrheitsfunktionalen Interpretation der Junktoren möglich scheint. In richtig gewählten Anwendungen kann Poole-Folgerung auf relativ einfache Weise Resultate erzeugen, die sonst nur mit erheblich höherem Aufwand möglich sind.

3.3 Abgeschirmte Folgerung. Die Bedingungen in Poole-Folgerungen unterscheiden sich von den Defaults nicht nur darin, daß sie eine rein negative Wirkung haben; sie sind zudem auch *abgeschirmt* gegen die Prämissen. Damit ist gemeint: Prämissen können zwar Defaults ausschalten, sie können aber keine Bedingungen ausschalten. Wenn wir zu den Defaults eine abgeschirmte Menge von Bedingungen hinzunehmen, denen aber ebenfalls eine positive Rolle bei der Erzeugung von Folgerungen zuweisen, dann folgern wir unter Rückgriff auf *anfechtbare* und *unanfechtbare*, d.h. gegen die Prämissen abgeschirmten Defaults.[8] Eine solche Art des Folgerns findet sich unter dem Namen "screened consequence" (abgeschirmte Folgerung) in Makinson [198] (siehe auch [200, pp. 43f.]).

Definition 12. Eine Formel A ist eine *abgeschirmte Folgerung* aus X unter anfechtbaren Defaults K und abgeschirmten Defaults H ($X \mid\!\sim_{KH} A$) gdw $K', X \vdash A$, für alle $K' \in K_X$ so, daß $H \subseteq K'$. D.h.

$$\mathrm{C}_{KH}(X) = \bigcap\{\mathrm{Cn}(K' \cup X) : K' \in K_X \text{ und } H \subseteq K'\}.$$

[8] Die "Anfechtbarkeit" eines Defaults ist natürlich etwas anderes als die Anfechtbarkeit eines Schlusses.

Wir können uns abgeschirmte Folgerung als einen Fall der Folgerung nach Teilschnitt vorstellen. Unter den Mengen in K_X wählen wir nur diejenigen, die alle abgeschirmten, d.h. nicht anfechtbaren Defaults enthalten. Die Auswahlfunktion läßt sich also durch Rückgriff auf H sehr einfach definieren,

$$s_{KH}(K_X) = \{K' \in K_X : K' \subseteq H\},$$

woraufhin

$$\mathrm{C}_{KH}(X) = \bigcap\{\mathrm{Cn}(K' \cup X) : K' \in s_{KH}(K_X)\}.$$

Hier, wie auch in den oben besprochenen beiden anderen Theorien, haben wir ein Beispiel vor uns, wie die in der allgemeinen Theorie des Folgerns nach Teilsschnitt bloß postulierte Auswahl konkret realisiert werden kann.

Abgeschirmte Folgerung ist das einfachste Beispiel einer Folgerung unter Rückgriff auf *geschichtete* Annahmenmengen. Hier bestehen die Annahmen aus nur zwei Schichten: Einer Menge H von Defaults, die gegen Anfechtung durch die Prämissen vollständig abgeschirmt ist, und einer Menge K anfechtbarer Defaults. Die Einteilung der Defaults in solche, die ganz und solche, die gar nicht anfechtbar sind, ist der erste Schritt zur Modellierung der Vorstellung, daß Defaults mehr oder weniger anfechtbar sein können. Einige Modellierungsmöglichkeiten sind in [200, Kap. 2.3] skizziert. Sie reichen von einer linearen Anordnung von Teilmengen der Defaultmenge (wie in [38]), die dann nacheinander und kumulativ auf Konsistenz mit den Prämissen geprüft werden, bis zur Annahme einer Anfechtbarkeitsordnung auf der gesamten Formelmenge (wie in [102]).

4. Materiale Interpretationen

Im ersten Abschnitt über Folgerung unter materialen Annahmen K hatten wir die unvorsichtige, monotone Relation so definiert:

$$(*) \qquad X \vdash_K A \text{ gdw } \forall I \text{ mit } I(K) = 1: \text{ wenn } I(X) = 1, \text{ dann } I(A) = 1.$$

Das ist im wesentlichen die klassische semantische Folgerungsbeziehung mit Quantifikation nicht über *alle* Interpretionen (Modelle), sondern nur über solche, in denen die Annahmen K wahr sind – oder, wie wir auch sagen können: Folgerung in einem Bereich “materialer” Interpretationen. Im weiteren haben wir dann verschiedene Strategien kennengelernt, mit den fixierten Annahmen vorsichtig umzugehen. In der einen oder anderen Weise erwiesen sich diese als Fälle (oder eng verwandt mit) der Teilschnittoperation nach Definition 8:

$$C_K(X) = \bigcap\{\mathrm{Cn}(K' \cup X) : K' \in s_K(K_X)\},$$

für eine Menge $s_K(K_X)$ ausgewählter Elemente in der Familie K_X mit X maximal konsistenter Teilmengen von K; bzw. in relationaler Notation:

$$X \mathrel{|\!\sim}_K A \text{ gdw } \forall K' : \text{ wenn } K' \in s_K(K_X), \text{ dann } K', X \vdash A.$$

Diese Definition verschiebt den Fokus von bestimmten Interpretationen auf bestimmte Formelmengen. Zwar können wir im Prinzip zwischen beiden hin und her übersetzen. Aber der natürliche Übergang von unvorsichtig zu vorsichtig materialem Schließen im Anschluß an $(*)$ ist eigentlich ein anderer. Wir wollen die Hintergrundannahmen K vorsichtig verwenden, um aus Prämissen X zu schließen. Ausgehend von $(*)$ bedeutet das, daß wir nicht *alle* Interpretationen mit $I(K) = 1$ (K-Interpretationen) betrachten sollten, sondern nur solche, die im Hinblick auf X gefahrlos infrage kommen. Das legt nun die folgende Modellierung nahe.

Wir gehen aus von der Menge *Int* der Interpretationen $I : \mathrm{ATM} \longrightarrow \{0,1\}$ und betrachten eine Teilmenge κ daraus (κ-Interpretationen). Für eine Formelmenge X notieren wir die Menge der κ-Interpretationen, die X erfüllen mit $[\![X]\!]_\kappa$; sinngemäß gleich: $[\![A]\!]_\kappa$ für eine Formel A. Die Notwendigkeit einer Auswahl in einer Menge $[\![X]\!]_\kappa$ von Interpretationen modellieren wir mit einer Relation $<$, die irreflexiv und transitiv ist, also so etwas wie Präferenz ausdrücken kann. Gegeben $(\kappa, <)$ definieren wir eine Bestenauswahl (minimale Elemente unter $<$) in einer Menge x von κ-Interpretationen: Für alle $x \subseteq \kappa \subseteq Int$,

$$\begin{aligned} I \in \min\nolimits_<(x) \text{ gdw } & I \in x \text{ und} \\ & \forall I' \in \kappa : I' \not< I. \end{aligned}$$

Schließlich definieren wir die vorsichtig materiale Folgerungsrelation so:[9]

$$(\mathrel{|\!\sim}_{\kappa,<}) \qquad X \mathrel{|\!\sim}_{\kappa,<} A \text{ gdw } \min\nolimits_<([\![X]\!]_\kappa) \subseteq [\![A]\!]_\kappa.$$

Es ist klar, daß wir mit den κ-Interpretationen auf diejenigen Interpretationen zielen, welche die Hintergrundannahmen K (vgl. $(*)$ oben) wahr

[9] Das ist im wesentlichen die Grunddefinition für "*preferential models*", wie sie zuerst in Shoham [266] vorgestellt wurde. Vgl. auch die allgemeinere Definition in [197]

machen. Aber unsere jetzige Definition ist tatsächlich etwas allgemeiner. Sie bringt die folgende Variante von $(*)$ in einen vorsichtigen Modus:

$$(**) \qquad X \vdash_\kappa A \text{ gdw } \forall I \text{ mit } I \in \kappa: \text{ wenn } I(X) = 1, \text{ dann } I(A) = 1$$

– oder, rechts etwas kürzer wie in $(\mid\!\sim_{\kappa,<})$: gdw $[\![X]\!]_\kappa \subseteq [\![A]\!]_\kappa$. Wir wollen die Interpretationsmenge κ *definierbar* nennen, falls es eine Formelmenge K gibt so, daß

$$I \in \kappa \text{ gdw } I(K) = 1.$$

Ist κ definierbar durch eine Formelmenge K, dann sind $(*)$ und $(**)$ äquivalent. Wenn wir es mit Interpretationen über endlich vielen atomaren Formeln zu tun haben, dann ist jede Teilmenge solcher Interpretationen definierbar. Im allgemeinen, d.h. im Falle von Interpretationen über beliebigen (abzählbaren) Mengen von Atomen ist das jedoch nicht so: Nicht jede Menge von Interpretationen ist definierbar. Ein notwendiges und hinreichendes Kriterium für die Definierbarkeit von κ ist – so werden wir gleich sehen – die Kompaktheit von $\vdash_\kappa$.

Im Gegensatz zu $\vdash_K$, ist die in $(**)$ definierte Relation $\vdash_\kappa$ nicht allgemein kompakt: Es mag $X \vdash_\kappa A$ gelten, ohne daß es eine endliche Teilmenge X' von X gibt mit $X' \vdash_\kappa A$. Man betrachte beispielsweise die Menge κ aller Interpretationen über einer (abzählbar) unendlichen Menge X von Atomen – außer der Interpretation T, die *alle* Atome in X wahr macht. Da keine Interpretation in κ die Menge X erfüllt, gilt $X \vdash_\kappa A$ trivialerweise für beliebige Formeln A, also auch für $A = \bot$. Aber jede Menge $X' \subset X$ (und also auch jede endliche Teilmenge von X) hat die Eigenschaft, daß es eine Interpretation in κ gibt, unter der X' wahr und die Kontradiktion $\bot$, natürlich, falsch ist; also gibt es insbesondere keine endliche Teilmenge X' von X mit $X' \vdash_\kappa \bot$.

Satz 13. *Die Relation $\vdash_{\kappa,<}$ ist genau dann kompakt, wenn κ definierbar ist.*

Beweis. $(\Leftarrow)$ Wenn κ definierbar ist, dann gibt es eine Menge K mit $I \in \kappa$ gdw $I(K) = 1$. Aus der Definition $(*)$ ergibt sich so, daß $\vdash_\kappa = \vdash_K$. Von der Relation $\vdash_K$ wissen wir, daß sie kompakt ist.

$(\Rightarrow)$ Die Relation $\vdash_{\kappa,<}$ ist supraklassich, reflexiv, monoton, und disjunktiv (in den Prämissen). Wenn wir jetzt noch annehmen, $\vdash_{\kappa,<}$ sei kompakt, dann handelt es sich um eine monoton materiale Folgerungsrelation im Sinne von Satz 3. Aber dann ist (nach Satz 3) κ definierbar. ■

Als Korollar folgt aus dem Satz, daß in unserem obigen Beispiel die Menge *Int* $\setminus$ T aller Interpretationen (über einer unendlichen Menge ATM von

Atomen) ausschließlich der Wahrmacherinterpretation T nicht definierbar ist in der aus ATM generierten Sprache.

Zurück zur vorsichtig materialen Relation $\mid\!\sim_{\kappa,<}$. Auch sie ist nicht allgemein kompakt. Das folgt schon aus dem Umstand, daß die unvorsichtige Relation aus $(**)$ sich erwartungsgemäß als Spezialfall der vorsichtigen darstellt. Wenn nämlich in $(\kappa, <)$ die Ordnung $<$ leer ist (keine Hintergrundannahme aus Vorsicht zurückgestellt wird), dann sind alle Interpretationen in κ minimal unter $<$, d.h. $\min_<(\llbracket X \rrbracket_\kappa) = \llbracket X \rrbracket_\kappa$. Ansonsten hat die Relation $\mid\!\sim_{\kappa,<}$ aber beinahe alle Eigenschaften, die wir schon für die Relation $\mid\!\sim_K$ feststellen konnten. Sie ist (vgl. S. 423)

supraklassisch,
reflexiv,
kumulativ transitiv (Schnitt), und
disjunktiv (in den Prämissen).

Auffallend abwesend in der Liste ist die Eigenschaft der Vorsichtigen Monotonie,

$$\frac{X \mid\!\sim_{\kappa,<} A \qquad X \mid\!\sim_{\kappa,<} B}{X, A \mid\!\sim_{\kappa,<} B}.$$

Ein einfaches Gegenbeispiel entnehmen wir [197, S. 73]: Es sei κ die Menge $\{I_i : i \in \mathbf{N}\}$, also eine abzählbar unendliche Menge von Interpretationen, und $<$ sei die Umkehrung der natürlichen Ordnung auf $\mathbf{N}$, also $... < I_1 < I_0$. Wir betrachten drei Atome P, Q, R mit $I(P) = 0$ und $I(Q) = 1$ für alle $I \in \kappa$, und für R sei nur $I_0(R) = 1$. Dann ist $\min_<(\llbracket P \rrbracket_\kappa) = \emptyset$ und also $P \mid\!\sim_{\kappa,<} Q$ und $P \mid\!\sim_{\kappa,<} R$. Da $I_0(Q) = 1 = I_0(R)$, so ist $\min_<(\llbracket Q, R \rrbracket_\kappa) = \{I_0\}$. Jedoch ist $I_0(P) = 0$. Also $Q, R \not\mid\!\sim_{\kappa,<} P$.

Problematisch für Vorsichtige Monotonie sind offenbar unendlich absteigende Ketten, welche die Prämissen auf leere Weise wahr machen können. Der Schluß ist nur gültig in Strukturen $(\kappa, <)$, in denen solche Ketten nicht auftreten, d.h. in denen es für jede in κ erfüllbare Formelmenge minimale Interpretationen gibt, d.h.

wenn $\llbracket X \rrbracket_\kappa \neq \emptyset$, dann ist $\min(X) \neq 0$.

In der Literatur werden $(\kappa, <)$-Strukturen mit dieser Eigenschaft *stoppered* [197] oder auch *smooth* [163] genannt. Endliche $(\kappa, <)$-Strukturen erfüllen die Bedingung *per se*. Die Eigenschaft ist ebenfalls hinreichend für den Schluß der Konsistenzbewahrung: Wenn $X \nvdash \bot$, dann auch $X \not\mid\!\sim_{\kappa,<} \bot$.

Die *Stopperedness*-Bedingung, sowie die Definition $(\mid\!\sim_{\kappa,<})$ überhaupt erinnert natürlich sehr an Lewis' Sphären-Semantik kontrafaktischer Konditionale. In Sphärenmodellen, so haben wir gesagt, ist das Konditional

$A \sqsupset B$ an einem Punkt a genau dann wahr, wenn die (von a aus gesehen) kleinste (minimale) $[\![A]\!]$-Sphäre in $[\![B]\!]$ enthalten ist. Das wollen wir jetzt (ein wenig anders als im Kapitel IV über Konditonale) kurz so notieren:

$$(\sqsupset) \qquad a \models A \sqsupset B \text{ gdw } \min_a([\![A]\!]) \subseteq [\![B]\!],$$

und sehen sogleich die starke Analogie zu ($\mid\!\sim_{\kappa,<}$). Um auch in unendlich verschachtelten Sphärensystemen zu garantieren, daß die Funktion min immer definiert ist, wurde dort die Limes-Annahme eingeführt. Sie spielt offensichtlich eine der *Stopperedness*-Bedingung völlig analoge Rolle. Um den Vergleich zwischen $\mid\!\sim$ und $\sqsupset$ weiter zu präzisieren, schränken wir einerseits die linke Seite von $\mid\!\sim$ so ein, daß dort nur Formeln stehen können, $\mid\!\sim$ also zu einer Relation FML $\times$ FML wird. Andererseits betrachten wir nur das sogenannte flache Fragment einer Sprache mit Konditionalen, d.h. $\sqsupset$ kann nur rein Boole'sche Formeln binden. Dadurch verliert $\sqsupset$ den Charakter eines (iterierbaren) Operators und kann, wie $\mid\!\sim$, als eine Relation zwischen (Boole'schen) Formeln aufgefaßt werden. Nach diesen syntaktischen Manövern können wir Postulate für $\mid\!\sim$ mit solchen für $\sqsupset$ direkt vergleichen. Auf der semantischen Seite fassen wir die Punkte eines Sphärenmodells als Interpretationen auf, bzw. bilden, umgekehrt, $(\kappa, <)$-Strukturen zu Modellen um, deren Punktmengen die Interpretationen in κ indizieren. Damit sei hier nur angedeutet, wie man zu systematischen Vergleichen zwischen konditionaler und nichtmonotoner Logik vordringen kann; vgl. z.B. [15, 16].

5. Materiale Regeln

Logische Regeln können wir als Paare von Formeln, (A, B) auffassen. Mehrprämissige Regeln,

$$\frac{A_1, ..., A_n}{B}$$

stellen wir durch Konjunktion der Prämissen dar: $(A_1 \wedge \cdots \wedge A_n, B)$. Im folgenden werden wir auch $A \Rightarrow B$ für eine Regel (A, B) schreiben. Wenn R eine Menge von Regeln ist, dann soll $R(X)$ das *Bild* der Formelmenge X unter den Regeln in R sein:

$$B \in R(X) \text{ gdw } \exists A \in X : A \Rightarrow B \in R.$$

Eine Formelmenge X ist unter einer Regelmenge R *abgeschlossen*, wenn $R(X) \subseteq X$.

Einige Erläuterungen sind hier nötig. Erstens, läßt sich leicht zeigen, daß R idempotent, $R(R(X) \subseteq R(X)$, und monoton ist, d.h. $R(X) \subseteq R(X')$, falls $X \subseteq X'$. Die Inklusionsbedingung $X \subseteq R(X)$ ist jedoch nicht allgemein erfüllt. Daher ist die Abbildung R keine eigentliche Abschlußoperation. (Wir könnten Inklusion erzwingen, indem wir das Definiens für $R(X)$ um "oder $B \in X$" erweitern oder fordern, daß zu einer Regelmenge immer alle Regeln der Form $A \Rightarrow A$ gehören.)

Zweitens, dürfen wir hier Regeln nicht im schematischen Sinne auffassen. In der Angabe (A, B) einer Regel stehen A und B für konkrete Formeln – A und B sind nicht Platzhalter für beliebige Formeln einer bestimmten Form. In der eingangs eingeführten Terminologie betrachten wir hier also *materiale*, nicht *formale* Regeln. Würden wir Regeln formal (schematisch) verstehen, dann könnten wir sie nicht im hier beabsichtigten, die logischen Regeln erweiternden Sinne verwenden. Denn aufgrund der deduktiven Vollständigkeit der klassischen Logik würde die Hinzunahme jeder schematischen Regel, die klassisch nicht zulässig ist, das Schließen trivialisieren; vgl. dazu Satz 2 oben.

Wir wollen nun materiale Regeln R ähnlich wie zuvor materiale Annahmen K nutzen. D.h. wir wollen eine gegebene Menge von Prämissen nicht nur unter die formalen Regeln der Logik stellen, sondern auch unter besondere materiale Regeln, um aus den Prämissen mehr als bloß logisch darin enthaltene Information zu gewinnen. Auf den ersten Blick scheint es spitzfindig und müßig zu sein, zwischen dem Schließen unter materialen Annahmen und dem unter materialen Regeln zu unterscheiden. Können wir die Regeln nicht als konditionale Annahmen formulieren? Und können wir die Annahmen nicht als Konklusionen aus einer trivial gegebenen Prämisse, $\top$, betrachten? Die Antwort auf die zweite Frage lautet Ja. Die Simulation von Regeln durch Konditionale funktioniert aber nicht uneingeschränkt, wie wir gleich sehen werden. Daher ist das Folgern unter Regeln eine echte Verallgemeinerung des Folgerns unter Annahmen.

Eine zunächst naheliegende Möglichkeit, materiale Regeln zu nutzen, besteht darin, die gegebenen Prämissen X unter den Regeln R abzuschließen und dann aus dem Bild $R(X)$ zusammen mit X zu schließen. In diesem Fall betrachten wir also $\mathrm{Cn}(R(X) \cup X)$. (Die Hinzunahme von X zu $\mathrm{Cn}(X)$ ist nötig, da – wie oben bemerkt – $X \subseteq \mathrm{Cn}(X)$ nicht garantiert ist.) Aber diese Möglichkeit kann nicht befriedigen, denn $\mathrm{Cn}(R(X) \cup X)$ muß selbst nicht unter R abgeschlossen sein, in welchem Fall wir nicht den gewünschten vollständigen Gebrauch von R machen würden.

Beispiel: Es sei $R = \{P \Rightarrow Q, P \vee R \Rightarrow S\}$. Wenn wir jetzt unter Rückgriff auf R die erweiterten Konsequenzen aus der Prämissenmenge $\{P\}$ bestimmen wollten, würden wir darunter auch S erwarten. Denn aus P folgt Q nach R; aus P folgt $P \vee Q$; schließlich folgt S aus $P \vee Q$ nach R. Aber $S \notin \mathrm{Cn}(R(\{P\})) = \mathrm{Cn}(\{P, Q\})$.

Die folgende Definition (*pivotal consequence* in [200]) beseitigt diesen Fehler.

Definition 14. (Monotones Folgern unter materialen Regeln.) Eine Formel A folgt aus X unter materialen Regeln R (Notation: $X \vdash_R A$ bzw. $A \in \mathrm{Cn}_R(X)$) gdw A in allen Erweiterungen von X ist, die unter Cn und R abgeschlossen sind, d.h.,

$$\mathrm{Cn}_R(X) = \bigcap\{X' \supseteq X : \mathrm{Cn}(X') \subseteq X' \text{ und } R(X') \subseteq X'\}.$$

Die Inklusionsbedingung $X \subseteq \mathrm{Cn}_R(X)$ ist nun aufgrund der Bedingung $X \subseteq X'$ erfüllt. Der Schnitt der Mengen X' im Definiens ist nicht leer, da trivialerweise die Menge aller Formeln den Bedingungen für X' genügt. Ferner läßt sich leicht zeigen, daß $\mathrm{Cn}_R(X)$ nun wie gewünscht sowohl unter Cn als auch unter R abgeschlossen ist.

Das Schließen unter materialen Regeln subsumiert das Schließen unter materialen Annahmen, im Sinne der Definition von $\vdash_K$ auf S. 414. Denn eine Annahme A können wir auch äquivalent als eine Regel aus der *Verum*-Prämisse, d.h. als $(\top, A)$ auffassen. So können wir eine Menge K von Annahmen immer umformen in eine Menge $\mathrm{reg}(K)$ von Regeln (der "Regularisierung" von K):

$$\mathrm{reg}(K) = \{\top \Rightarrow A : A \in K\}.$$

Es läßt sich dann leicht zeigen, daß sich das Folgern unter materialen Annahmen immer als Folgern unter materialen Regeln darstellen läßt, nämlich so:

$$\mathrm{Cn}_K = \mathrm{Cn}_{\mathrm{reg}(K)}.$$

Umgekehrt aber läßt sich das Folgern unter Regeln im allgemeinen nicht als Folgern unter Annahmen darstellen. Das werden wir gleich mit einem Beispiel illustrieren. Unter welcher Bedingung die Simulation von Reglen durch Annahmen gelingt, werden wir in den Sätzen 17 und 18 festhalten.

Die Definition 14 fordert im Definiens den simultanen Abschluß unter Cn und unter R. Es ist nicht unmittelbar klar, wie das in eine Konstruktion des Definiendums umzusetzen ist. Naheliegend ist jedoch dieses Verfahren: Wir beginnen mit $\mathrm{Cn}(X)$. Dann schließen wir $X \cup R(X)$ unter Cn ab. Das Resultat, $X_1 = \mathrm{Cn}(X \cup R(X))$, schließen wir wieder unter R ab, worauf wir $X_1 \cup R(X_1)$ wiederum unter Cn abschließen. Mit dieser neuen Mengen X_2 verfahren wir genauso, u.s.w. Das folgende Lemma hält fest, daß dieses induktive Verfahren tatsächlich $Cn_R(X)$ konstruiert.

LEMMA 15. (Makinson) $\text{Cn}_R(X) = \bigcup X_n\,(0 \leq n < \omega)$, *wobei* $X_0 = \text{Cn}(X)$ *und* $X_{n+1} = \text{Cn}(X_n \cup R(X_n))$.

BEWEIS. (Wir kürzen $\bigcup X_n\,(0 \leq n < \omega)$ zu α ab.)

$\subseteq$: Wir beweisen die Behauptungen 1–3: α erweitert X und ist unter Cn und R abgeschlossen. Dann schließen wir nach Definition 14, daß α eine der im Definiens für $\text{Cn}_R(X)$ zu schneidenden Mengen ist, woraus folgt, daß der Schnitt, $\text{Cn}_R(X)$, in α enthalten sein muß.

1. $X \subseteq \alpha$ folgt unmittelbar aus der Basis der Definition von α.
2. $\text{Cn}(\alpha) \subseteq \alpha$. Man nehme an, daß $\alpha \vdash A$. (Wir zeigen, daß $A \in \alpha$.) Dann (Kompaktheit!) gibt es eine Teilmenge $X' \subseteq \alpha$ so, daß $X' \vdash A$ und $X' \subseteq X_n \subseteq \alpha$, für eine der induktiv definierten Mengen X_n; also auch $X_n \vdash A$. Da aber $X_n \subseteq \text{Cn}(X_n)$, so ist $A \in X_n \subseteq \alpha$.
3. $R(\alpha) \subseteq \alpha$. Angenommen $A \in \alpha$ und $A \Rightarrow B$. Dann gibt es eine der induktiv definierten Mengen $X_n \subseteq \alpha$ mit $\alpha \in X_n$. Da $X_n \subseteq R(X_n)$, so ist $B \in X_n \subseteq \alpha$, wie gewünscht.

$\supseteq$: Durch Induktion über n zeigen wir, daß jede der Mengen der im Definiens von α vereinigten Mengen X_n in $\text{Cn}_R(X)$ enthalten ist. Es folgt dann, daß die Vereinigung dieser Mengen, α, in $\text{Cn}_R(X)$ enthalten ist.

Für X_0 gilt die Behauptung, da $\vdash\, \leq\, \vdash_R$.

Für X_{n+1} nehmen wir an, daß $A \in X_{n+1}$, d.h.

$$A \in \text{Cn}(X_{n+1} \cup R(X_{n+1})) \tag{1}$$

und fahren fort unter der Induktionsannahme

$$\text{Cn}(X_n \cup R(X_n)) \subseteq \text{Cn}_R(X). \tag*{IA.}$$

(Zu zeigen: $A \in \text{Cn}_R(X)$.) (1) ist nach Def. äquivalent zu

$$A \in \text{Cn}(\,\text{Cn}(X_n \cup R(X_n)) \cup R(\text{Cn}(X_n \cup R(X_n)))\,). \tag{2}$$

Aufgrund der IA folgt daraus

$$A \in \text{Cn}(\,\text{Cn}_R(X) \cup R(\text{Cn}_R(X))\,). \tag{3}$$

Nun ist aber Cn_R unter R abgeschlossen, weshalb $R(\text{Cn}_R(X)) = \text{Cn}_R(X)$. Also folgt aus (3),

$$A \in \text{Cn}(\text{Cn}_R(X))$$

Und da Cn_R auch unter Cn abgeschlossen ist, so erhalten wir die gewünschte Konklusion $A \in \text{Cn}_R(X)$. ∎

Die Induktion in Lemma 15 läßt sich auch “feinkörniger” darstellen, nämlich so, daß die Induktion in jedem Schritt durch die Anwendung genau einer Regel fortschreitet. Dazu ordnen wir die Regeln R nach dem Ordnungstyp der natürlichen Zahlen durch eine Relation $\prec$ an und bestimmen die Konsequenzen $\mathrm{Cn}_{R\prec}(A)$ einer Menge A unter den so angeordneten Regeln R. Es sei:

$$X_0 = \mathrm{Cn}(X).$$

$$X_{n+1} = \begin{cases} \mathrm{Cn}(X_n \cup \{B\}), \text{ falls } A \Rightarrow B \text{ die erste Regel} \\ \quad \text{in der Ordnung ist, so daß } A \in X_n \text{ und } B \notin X_n; \\ X_n, \text{ falls es keine solche Regel gibt.} \end{cases}$$

$$\mathrm{Cn}_{R\prec}(X) = \bigcup X_n \, (0 \leq n < \omega).$$

Dieses Muster einer induktiven Definition des Schließens unter Bezug auf eine Anordnung der Hintergrundregeln wird im nächsten Abschnitt eine wichtige Rolle spielen. Hier ist die Anordnung jedoch überflüssig.

KOROLLAR 16. *Für jede Anordnung $\prec$ der Regeln R, $\mathrm{Cn}_{R\prec}(A) = \mathrm{Cn}_R(A)$.*

BEWEIS. Aus dem Lemma. ∎

Einige Eigenschaften von Cn_R bzw. $\vdash_R$. Aus der Definition gehen einige Eigenschaften von $\vdash_R$ unmittelbar hervor. So ist Cn_R supraklassisch, reflexiv, monoton und hat die Schnitteigenschaft. Es läßt sich ebenfalls zeigen, daß die regelbasierte Folgerung $\vdash_R$ wie die annahmenbasierte $\vdash_K$ kompakt ist.

Im Gegensatz zur annahmenbasierten erfüllt die regelbasierte Relation jedoch nicht ohne weiteres die Bedingung der Disjunktivität in den Prämissen,

Disjunktiv $$\frac{X, A \vdash_R C \qquad X, B \vdash_R C}{X, A \vee B \vdash_R C}.$$

Gegenbeispiel zu Disjunktiv: Es sei $R = \{P \Rightarrow S, Q \Rightarrow S\}$, für distinkte Atome P, Q, S. Dann $P \vdash_R S$ und $Q \vdash_R S$. Da weder $P \vee Q \vdash P$ noch $P \vee Q \vdash Q$, so können die Regeln in R durch $P \vee Q$ nicht ausgelöst werden. Daher ist $R(\mathrm{Cn}(P \vee Q)) = \mathrm{Cn}(P \vee Q)$. Da aber nicht $P \vee Q \vdash S$, so gilt nach der Definition daher auch nicht $P \vee Q \vdash_R S$.

Da im Gegensatz zur regelbasierten die annahmenbasierte Folgerung prämissendisjunktiv ist, zeigt diese Beobachtung, daß es einen wesentlichen Unterschied macht, ob wir unsere Prämissen unter einer Regel $A \Rightarrow B$

abschließen oder um eine Annahme $A \to B$ erweitern. Die Ersetzung von Regeln durch ein entsprechendes materiales Konditional wollen wir die “Materialisierung” von Regeln nennen. Wenn R eine Regelmenge ist, dann sei die Materialisierung $\mathrm{mat}(R)$ von R so defniert:

$$\mathrm{mat}(R) = \{a \to c : a \Rightarrow c \in R\}.$$

Das gerade angegebene Gegenbeispiel zur Disjunktivität von $\vdash_R$ können wir auch nutzen, um an einem Beispiel zu belegen, daß Regeln sich nicht durch ihre Materialisierungen simulieren lassen: $\mathrm{Cn}_R \neq \mathrm{Cn}_{\mathrm{mat}(R)}$.

> Es sei R wie im Beispiel oben und also $\mathrm{mat}(R) = \{P \to S, Q \to S\}$; daher wie zuvor $S \notin \mathrm{Cn}_R(P \vee Q)$. Nach der Definition des Schließens unter Annahmen haben wir $B \in \mathrm{Cn}_{\mathrm{mat}(R)}(A)$ gdw $A, \mathrm{mat}(R) \vdash B$. Also $P, \mathrm{mat}(R) \vdash S$ und $Q, \mathrm{mat}(R) \vdash S$. Daraus folgt nach *Disjunktiv* für $\vdash$, daß $P \vee Q, \mathrm{mat}(R) \vdash S$, d.h. $S \in \mathrm{Cn}_{\mathrm{mat}(R)}(P \vee Q)$.

Eine Hälfte der Identität $\mathrm{Cn}_{\mathrm{mat}(R)} = \mathrm{Cn}_R$, nämlich $\mathrm{Cn}_R \leq \mathrm{Cn}_{\mathrm{mat}(R)}$, wird durch das Gegenbeispiel nicht berührt. Daß diese Inklusion tatsächlich ohne Einschränkung gilt, können wir dem ersten Teil des Beweises von Satz 18 (s.u.) entnehmen.

Der Umstand, daß *Disjunkiv* für die Relation $\vdash_R$ nicht gilt, spricht nicht gerade für diese Relation im Sinne einer Folgerungsrelation. Denn einerseits dürfen wir von Folgerungsrelationen erwarten, daß *Disjunkiv* ausnahmslos gilt. Andererseits ist es so, daß wenn wir den Spielraum nutzen, die Regelmenge R so zu wählen, daß $\vdash_R$ die in *Disjunkiv* beschriebene Eigenschaft hat, dann läßt sich $\vdash_R$ immer auf eine annahmenbasierte Relation $\vdash_K$ reduzieren. Das ist der Inhalt des folgenden Satzes, demzufolge es genau die Eigenschaft *Disjunkiv* ist, welche den Unterschied markiert zwischen dem Schließen unter Rückgriff auf Annahmen und dem unter Rückgriff auf Regeln.

SATZ 17. *Eine regelbasierte Operation* Cn_R *ist genau dann als annahmenbasierte Operation* Cn_K *darstellbar, wenn jene disjunktiv in den Prämissen ist.*

BEWEIS. Wir wissen bereits, daß jede annahmenbasierte Operation Cn_K als regelbasiert, nämlich durch $\mathrm{Cn}_{\mathrm{reg}(K)}$ darstellbar ist. — Umgekehrt ist jede regelbasierte Operation Cn_R supraklassisch, reflexiv, monoton und kompakt. Unter der Annahme, daß Cn_R auch *Disjunktiv* erfüllt, ist Cn_R eine monoton materiale Folgerungsoperation. Wir können also Satz 3 anwenden und schließen, daß es eine Menge K von Annahmen gibt, so daß $\mathrm{Cn}_R = \mathrm{Cn}_K$. ■

Der nächste Satz identifiziert – nicht überraschend – die Annahmenmenge K für die Cn_R mit Cn_K übereinstimmt, falls Cn_R die Bedingung *Disjunkiv* erfüllt.

Satz 18. (Vgl. Makinson und van der Torre [207, Obs. 16] und Makinson [200, Thm. 4.4]). *Wenn* Cn_R *die Bedingung Disjunkiv erfüllt, dann ist* $\mathrm{Cn}_R = \mathrm{Cn}_{\mathrm{mat}(R)}$.

Beweis. Erinnerung:

$$(\alpha =) \quad \mathrm{Cn}_R(X) = \bigcap\{X' : X \subseteq X' \;\&\; \mathrm{Cn}(X') \subseteq X' \;\&\; R(X') \subseteq X'\};$$
$$(\beta =) \quad \mathrm{Cn}_{\mathrm{mat}(R)}(X) = \mathrm{Cn}(X \cup \mathrm{mat}(R)).$$

$\alpha \subseteq \beta$, d.h. $\mathrm{Cn}_R \leq \mathrm{Cn}_{\mathrm{mat}(R)}$: Es ist klar, daß $X \subseteq \beta$ und daß $\mathrm{Cn}(\beta) \subseteq \beta$. Ebenso ist $R(\beta) \subseteq \beta$. Denn, angenommen $A \in \beta$ und $A \Rightarrow B \in R$. Dann ist $A \to B \in \mathrm{mat}(R) \subseteq \beta$ und somit $B \in \beta$, da β unter Cn abgeschlossen ist. Also ist β eine der Mengen X' im Definiens von α, weshalb der Schnitt dieser Mengen in β enthalten sein muß.

Für die Umkehrung $\beta \subseteq \alpha$, d.h. $\mathrm{Cn}_{\mathrm{mat}(R)} \leq \mathrm{Cn}_R$ bieten wir zwei Beweise an. Der erste nimmt als Ausgangspunkt eine Beobachtung, die wir im Laufe des Beweises für unseren ersten Repräsentationssatz 3 machen konnten. Der zweite Beweis verfährt per Induktion über die Anzahl der Regeln in der Menge R. Man beachte in beiden Beweisen den Einsatz der Annahme, daß Cn_R unter *Disjunkiv* abgeschlossen ist.

Erster Beweis: Da Cn_R nach Annahme monoton material ist, können wir wie im Beweis von Satz 3 vorgehen und schließen, daß $\alpha = \mathrm{Cn}(A \cup \mathrm{Cn}_R(\emptyset))$. Es genügt nun zu zeigen, daß $\mathrm{mat}(R) \subseteq \mathrm{Cn}_R(\emptyset)$. Wir nehmen also an, daß (1) $A \Rightarrow B \in R$ und zeigen, daß $A \to B \in \mathrm{Cn}_R(\emptyset)$. Aus (1) folgt, daß $B \in \mathrm{Cn}_R(A)$ und da $\mathrm{Cn} \leq \mathrm{Cn}_R$ folgt ferner, (2) $A \to B \in \mathrm{Cn}_R(A)$. Ebenso haben wir (3) $A \to A \in \mathrm{Cn}_R(\neg A)$. Da wir annehmen, daß Cn_R unter *Disjunktiv* abgeschlossen ist, folgt aus (2) und (3), daß $A \to B \in \mathrm{Cn}_R(A \vee \neg A) = \mathrm{Cn}_R(\emptyset)$.

Zweiter Beweis: Da $\mathrm{C}_{\mathrm{mat}(R)}$ kompakt ist, gibt es eine endliche Teilmenge $M \subseteq \mathrm{mat}(R)$ so, daß $\mathrm{Cn}_{\mathrm{mat}(R)}(X) = \mathrm{Cn}_M(X)$. Wir nehmen nun an, daß $C \in \mathrm{Cn}_M(X)$, d.h. (1) $X, M \vdash C$, und zeigen durch Induktion über die Anzahl $k(M)$ der Implikationen in M, daß $C \in \alpha$. Im Falle $k(M) = 0$ ist $\alpha = \mathrm{Cn}(X) = \beta$. Der nicht-triviale Teil der Induktion beginnt daher mit $k(M) = 1$.

Basis $k(M) = 1$: Dann ist $M = \{A \to B\}$ und (1) wird zur Annahme (2) $X, A \to B \vdash C$. Da Cn_R unter Cn abgeschlossen ist, gilt $(*) \vdash \leq \vdash_R$. So folgt aus (2) nach $(*)$, (3) $X, A \to B \vdash_R C$. Ebenfalls nach $(*)$ haben wir

(4) $\neg A \vdash_R A \to B$. Aus (3) und (4) folgt durch Schnitt, (5) $X, \neg A \vdash_R C$. Da Cn_R unter der Regel $A \Rightarrow B$ abgeschlossen ist, haben wir $X, A \vdash_R B$ und also nach $B \vdash_R A \to B$ $(*)$!, (6) $X, A \vdash_R A \to B$. Aus (3) und (6) folgt durch Schnitt, (7) $X, A \vdash_R C$. Aus (5) und (7) folgt schließlich nach *Disjunktiv* (und Schnitt), daß $X \vdash_R C$.

Schritt: Wir machen nun die Induktionsannahme für $k(M) = n$, d.h. (1) $\mathrm{Cn}(X, M) \subseteq \mathrm{Cn}_{\mathrm{reg}(M)}(X)$, und zeigen, daß die Behauptung dann für $k(M) = n + 1$ gilt. Wir nehmen also ferner an, daß (2) $X, M, A \to B \vdash C$ und zeigen, daß $X \vdash_R C$ mit $R = \mathrm{reg}(M) \cup \{A \Rightarrow B\}$. Aus (2) folgt (3) $X, M \vdash (A \to B) \to C$ und so nach der I.A. (1), daß (4) $X \vdash_{\mathrm{reg}(M)} (A \to B) \to C$. Wir wissen, daß $(*)$ $\mathrm{Cn} \leq \mathrm{Cn}_{\mathrm{reg}(M)} \leq \mathrm{Cn}_R$. Also folgt aus (4), (5) $X, A \to B \vdash_R C$. Der Beweis zur Konklusion (6) $X, \neg A \vdash_R C$ ist nun wie im Basisfall. Ebenso erhalten wir die Konklusion (7) $X, A \vdash_R C$ wie oben. Aus (6) und (7) folgt dann $X \vdash_R C$ nach *Disjunktiv* wie gewünscht.
■

So, wie sich nach Satz 3 jede monoton materiale Folgerungsoperation Cn^+ als eine annahmenbasierte Operation Cn_K darstellen läßt, so läßt sich eine solche Operation Cn^+ übrigens auch als regelbasierte Operation Cn_R darstellen. Allerdings ist das, im Unterschied zu Satz 3, eine nicht sehr interessante Repräsentationsmöglichkeit. Sie beruht nämlich darauf, daß wir die zu Cn^+ passende Regelmenge immer so festlegen können: $A \Rightarrow B \in R$ gdw $A \vdash^+ B$ – ersichtlich eine triviale Repräsentation; vgl. [200, Thm 4.5].

6. Vorsichtige Anwendung materialer Regeln

Wie das annahmenbasierte Schließen, so ist auch das regelbasierte Schließen monoton:

(M) Wenn $X \subseteq Y$, dann $\mathrm{Cn}_R(X) \subseteq \mathrm{Cn}_R(Y)$.

Die Monotonie-Eigenschaft (M) reflektiert den Umstand, daß eine Prämissenmenge X immer alle anwendbaren Regeln in R auslöst. Je mehr Prämissen zur Verfügung stehen, desto mehr Regeln werden ausgelöst. Dies geschieht auch dann, wenn dadurch widersprüchliche Konklusionen erzeugt werden.

Beispiel: Es sei $R = \{P \Rightarrow S, Q \Rightarrow \neg S\}$. Dann $P, Q \vdash_R \bot$.

In dem Beispiel würden wir uns wünschen, daß unter den Prämissen $\{P, Q\}$ nicht beide Regeln zur Anwendung kommen. Die eine oder die andere der Regeln, vielleicht beide, sollte angesichts der Prämissen deaktiviert sein. Im allgemeinen möchten wir die Frage, welche Regeln in einer Regelmenge aktiv sein sollen und also durch gegebene Prämissen ausgelöst werden können, in Abhängigkeit von diesen Prämissen beantworten. Wie immer das im Detail aussehen mag, im Resultat wird die Monotonie-Eigenschaft (M) verloren gehen.

In groben Zügen gibt es zwei Strategien, das Konsistenzproblem für das Schließen aus Default-Regeln – wie wir jetzt sagen wollen – zu lösen. Wir können vorgehen wie im Falle des Schließens aus Default-Annahmen. Dann schneiden wir die Menge der anwendbaren Regeln auf maximale Mengen zurück, die sich konsistent auf die jeweiligen Prämissen anwenden lassen. Alternativ können wir eine Reihenfolge festlegen – oder mehrere solcher Reihenfolgen betrachten –, in der wir die Regeln anwenden wollen und arbeiten diese Reihe ab, solange dies mit den jeweiligen Prämissen zu konsistenten Ergebnissen führt.

Die erste Strategie verfährt für regelbasierte Operationen im wesentlichen nicht anders als für annahmenbasierte Operationen. Als Ausgangspunkt nehmen wir die Definition 14 der monotonen Operation Cn_R,

$$\mathrm{Cn}_R(X) = \bigcap\{X' \supseteq X : \mathrm{Cn}(X') \subseteq X' \ \& \ R(X') \subseteq X'\},$$

und ersetzen darin $R(X')$ durch das Bild von X' unter einer Regelmenge, die aus R und X nach Teilschnitt gebildet wird. Dafür bestimmen wir zunächst die maximalen Teilmengen von Regeln, die auf X konsistent angewandt werden können. Die Familie R_X dieser Teilmengen von R sei so definiert:

Def. R_X $$\begin{aligned} R' \in R_X \text{ gdw } R' \subseteq R,\ X \nvdash_{R'} \bot, \text{ und} \\ \forall S:\ \text{wenn } R' \subset S \subseteq R, \text{ dann } X \vdash_S \bot. \end{aligned}$$

Unter den Regelmengen in R_X wollen wir die *besten* auswählen, was wir wieder durch eine Auswahlfunktion s darstellen, welche wie zuvor unter der offensichtlichen Bedingung $s(R_X) \subseteq R_X$ steht.

Sodann fordern wir von den Mengen X' im Definiens von Cn_R nicht, daß sie unter *allen* Regeln in R abgeschlossen sind, sondern nur unter denen, die zu allen *besten* Regelmengen gehören. So ersetzen wir die Regelmenge R auf der rechten Seite der Definition 14 durch eine Konstruktion mithilfe eines Teilschnitts der Regeln, nämlich durch $\bigcap s(R_X)$ und fordern nun, daß X' unter dieser – typischerweise kleineren – Menge von Regeln abgeschlossen sein soll. Das Ergebnis ist diese Definition einer Operation des *Folgerns aus Default-Regeln nach dem Teilschnitt* (im Kontext stellt hier $[\bigcap s(R_X)]$ die Abschlußoperation dar):

Def. C_R $$\begin{aligned} \mathrm{C}_R(X) = \bigcap\{\ X' \subseteq X :& \mathrm{Cn}(X') \subseteq X' \\ & \text{und } [\textstyle\bigcap s(R_X)](X') \subseteq X'\ \}. \end{aligned}$$

Diese Definition subsumiert als Grenzfälle wieder (a) $s(R_X) = R_X$ und (b) $s(R_X) = \{R'\}$ für eine bestimmte Regelmenge $R' \in R_X$. Der Fall (a) ist ein Folgern nach dem vollen Schnitt; der Fall (b) beschreibt ein Folgern nach einer Maximalwahl.

Obgleich diese erste Strategie zur Herstellung von Konsistenz aufgrund ihrer strikten Analogie zum vorsichtigen annahmenbasierten Schließen völlig naheliegend ist, ist sie in der Literatur eher selten anzutreffen. Daher wollen wir es bei dieser Skizze bewenden lassen und wenden uns der zweiten Strategie zu, Reiters Default Logik.

7. Regeln unter Bedingungen

Wir beginnen damit, den Regelbegriff zu erweitern. Regeln seien nun Tripel $(A, \mathcal{B}, C)$. Die erste Koordinate, A, ist das Antezedens (Prämisse), die Formelmenge $\mathcal{B}$ enthält die Bedingungen und C ist das Konsequens (Konklusion) der Regel. Statt $(A, \mathcal{B}, C)$ schreiben wir auch $A : \mathcal{B} \Rightarrow C$. Regeln in diesem erweiterten Sinne werden wir *Defaults* nennen.[10] Ein Default $A : \mathcal{B} \Rightarrow C$ soll ausdrücken, daß wir von A auf C schließen können unter der Bedingung, daß das, was wir im Kontext des Schließens wissen können, mit $\mathcal{B}$ verträglich ist. (Wollen wir zum Beispiel aus P auf Q mit dem Default $P : \{R\} \Rightarrow Q$ schließen, dann müssen wir uns zunächst davon überzeugen, daß die Information $\neg R$ nicht vorliegt.)

Insofern die Regeln $A \Rightarrow C$ des letzten Abschnitts unter der Bedingung stehen, daß sie nur auf vorsichtige Weise angewendet werden sollen, können wir diese auch so verstehen: Schließe von A auf C unter der Bedingung, daß unsere Annahme mit C konsistent ist. Mit anderen Worten, vorsichtige Paar-Regeln (A, C) können wir als einen besonderen Fall von Tripel-Regeln auffassen, nämlich als Defaults der Form $(A, \mathcal{B}, C)$, mit $\mathcal{B}$ äquivalent zu $\{C\}$. (Im gerade angeführten Beispiel ersetze man R durch Q.) Vereinfachend wollen wir solche Regeln künftig (A, C, C) bzw. $A \dot{\Rightarrow} C$ schreiben. Solche Defaults heißen *normal*. Die jetzt zu besprechende Strategie ist also auch auf die einfacheren Regeln des vorigen Abschnitts anwendbar. Wir werden gleich darauf zurückkommen.

Wenn D eine Menge von Defaults und X eine Menge von Formeln ist, dann ist (D, X) eine *Default-Theorie*. Die Frage ist nun: Was dürfen wir aus einer Default-Theorie schließen? Oder, anders formuliert: Wie können wir die in X enthaltene Information mit Hilfe der Defaults D maximal und

[10] "Standardregel" wäre eine mögliche Übersetzung. Jedoch hat sich mittlerweile der englische Ausdruck "Default" im hier gemeinten, präzisen Sinne, der auf [243] zurückgeht, auch im Deutschen eingebürgert. In der Literatur werden die Bedingungen $\mathcal{B}$, ebenfalls im Anschluß an [243], auch "Rechtfertigungen" (*justifications*) genannt.

zugleich konsistent erweitern? Auch ohne eine formale Definition in der Hand, legen kleine Beispiele nahe, daß Default-Theorien im allgemeinen auf mehr als eine Weise im beabsichtigten Sinne erweitert werden können.

Beispiel: $X = \{P\}$, $D = \{P \dot{\Rightarrow} Q,\ P \dot{\Rightarrow} \neg Q\}$. Wenn wir zuerst X unter Rückgriff auf $P \dot{\Rightarrow} Q$ erweitern, so erhalten wir Q. Daraufhin ist eine Erweiterung nach $P \dot{\Rightarrow} \neg Q$ nicht mehr möglich. Wenn wir dagegen zuerst $P \dot{\Rightarrow} \neg Q$ auf X anwenden, dann erhalten wir eine andere Erweiterung. X läßt sich also auf zweierlei Weise erweitern.

Man beachte, daß wir jetzt die Frage nach Erweiterungen der Prämissenmenge X auf eine andere Weise stellen als in der oben besprochenen ersten Strategie. Dort haben wir zunächst nach maximalen Teilmengen der Regeln gefragt, die konsistent auf X angewandt werden können. Nun fragen wir direkt nach maximalen Erweiterungen von X, die konsistent aus X unter Verwendung der Regeln generiert werden können. Daß dies nicht nur verschiedene Perspektiven auf denselben Gegenstand sind, werden wir gleich sehen. Jedenfalls erfordert eine Antwort auf die gestellte Frage eine andere Art von Definition als in der ersten Strategie. Was wir jetzt benötigen, ist so etwas wie eine Induktion, die mit X beginnt und dann die Regeln sukzessive anwendet unter Beachtung der jeweiligen Konsistenzbedingungen. Wir werden drei solcher Definitionen vorstellen, welche formal recht verschieden, im Ergebnis aber äquivalent sind.

Definition 19. (Reiter [243]) Es sei (D, X) eine Default-Theorie. Man betrachte eine Folge von Formelmengen $X_0, X_1, X_2, ..., X^*$, die unter den folgenden Bedingungen steht:

$$\begin{aligned} X_0 &= X; \\ X_{n+1} &= X_n \cup \{C : (A, \mathcal{B}, C) \in D \text{ und } X_n \vdash A \text{ und } X^*, B \nvdash \bot\, (\forall B \in \mathcal{B})\}; \\ X^* &= \bigcup \{X_n : n < \omega\}. \end{aligned}$$

Wenn es eine solche Folge $X_0, X_1, X_2, ..., X^*$ gibt, dann ist $E = \mathrm{Cn}(X^*)$ eine *Reiter-Erweiterung* von (D, X).

Die Frage, was wir nun aus den möglicherweise vielen Reiter-Erweiterungen einer Default-Theorie schließen sollten, wollen wir erst im nächsten Abschnitt beantworten. Zunächst stellen wir fest, daß die Definition auf den ersten Blicke wie eine induktive aussieht. Bei genauerem Hinsehen wird jedoch im Schritt $n+1$ auf das Ergebnis der Konstruktion, X^*, *voraus*gegriffen. Die Definition verfährt also nicht induktiv im eigentlichen Sinne, sondern “quasi-induktiv”. Sie läßt sich äquivalent auch als Beschreibung einer Fixpunkt-Operation auffassen:

DEFINITION 20. (Reiter [243]) Es sei (D, X) eine Default-Theorie. Eine Formelmenge E ist genau dann eine *Fixpunkt-Erweiterung* von (D, X), wenn $\Gamma(E) = E$, wobei Γ für (D, X) und beliebige Formelmengen Y durch die folgenden Bedingungen definiert ist:

1. $X \subseteq \Gamma(Y)$;
2. $\mathrm{Cn}(\Gamma(Y)) \subseteq \Gamma(Y)$;
3. wenn $(A, \mathcal{B}, C) \in D$ und $\Gamma(Y) \vdash A$ und $Y, B \not\vdash \bot\, (\forall B \in \mathcal{B})$, dann $C \in \Gamma(Y)$;
4. $\Gamma(Y)$ ist die kleinste Menge, welche 1–3 erfüllt.

Wie schon angedeutet, beschreiben die quasi-induktive Charakterisierung in Definition 19 und die Fixpunkt-Gleichung in Definition 20 dieselben Formelmengen.

SATZ 21. *(Reiter) Eine Formelmenge E ist genau dann eine Reiter-Erweiterung von (D, X), wenn E eine Fixpunkt-Erweiterung von (D, X) ist.*

BEWEIS. Reiter [243, Thm 2.1]. Siehe auch Poole [220, Thm 4.1.1, pp. 199f] ∎

Die Funktion Γ in Definition 20 ist wohldefiniert: Für jede Default-Theorie und Formelmenge Y gibt es eine Menge $\Gamma(Y)$, welche die Bedingungen 1-4 der Definition erfüllt. Daraus sollte man nicht schließen, daß es für jede Default-Theorie auch eine Menge gibt, die ein Fixpunkt unter Γ ist. (Auch die Definition einer Reiter-Extension steht unter dem Vorbehalt, daß es eine quasi-induktive Folge von Formelmengen überhaupt gibt.) Tatsächlich können Default-Theorien eine, mehrere oder auch gar keine Erweiterungen haben.

Beispiele: Die Default-Theorie sei $(\{\top : P \Rightarrow \neg P\}, \emptyset)$. Es kann hier keine deduktiv abgeschlosse Menge Y geben, so daß $Y = \Gamma(Y)$ und die Bedingung 3 der Definition 20 erfüllt ist. Denn aufgrund der Annahmen folgt aus 3, daß $(*)$ wenn $Y, P \not\vdash \bot$, dann $\neg P \in Y$. Aber $(*)$ ist widersprüchlich: $Y, P \not\vdash \bot$ gdw $Y \not\vdash \neg P$ gdw (deduktiver Abschluß!) $\neg P \notin Y$.
Dieses minimale Beispiel läßt sich verallgemeinern. Eine Default-Theorie (D, X) sei derart, daß $A \in X$ und $(A : B \Rightarrow \neg B) \in D$. Ähnlich wie zuvor läßt sich dann zeigen, daß es keine deduktiv abgeschlosse Obermenge X' von X geben kann, welche zugleich die Bedingung 3 der Definition erfüllt und ein Fixpunkt unter Γ ist.

Die Definitionen bestimmen zwar, was eine Extension, d.h. eine maximal konsistente Anreicherung von Prämissen unter Rückgriff auf Defaults ist; sie tun dies aber auf eine etwas undurchsichtige, nicht-konstruktive Weise.

Können wir Erweiterungen auch durch eine richtige Induktion bestimmen? Eine solche Induktion würde uns eine Anweisung geben, wie wir Erweiterungen Schritt für Schritt aufbauen können, ohne dabei schon das Resultat voraussetzen zu müssen. Die Antwort auf die Frage fällt positiv aus. Für normale Defaults hat dies schon Brewka [40] gezeigt. Die hier dargestellte Lösung für den allgemeineren Fall stammt von Makinson [199] (siehe auch [200]).

Die Definition ist etwas komplexer als die bisher betrachteten, in ihrer Grundidee aber einfach.

> Wir wollen sagen, ein Default $(A, \mathcal{B}, C)$ sei auf eine Formelmenge X *anwendbar*, wenn $X \vdash A$ und $X, B \nvdash \bot$ $(\forall B \in \mathcal{B})$.

Zunächst legen wir eine Reihenfolge fest, in der die Defaults angewandt werden sollen. Wir beginnen die Konstruktion mit den deduktiven Konsequenzen der Ausgangsmenge von Prämissen und führen in jedem Schritt Buch darüber, welche Defaults bisher angewandt, "verbraucht" wurden. Zu Beginn ist die Menge der verbrauchten Defaults also leer. In jedem weiteren Schritt wird nach dem ersten noch nicht verbrauchten und in diesem Schritt anwendbaren Default gesucht. Das Resultat des Schritts ist einerseits die nach der Anwendung des Defaults erreichte Erweiterung der Prämissenmenge, andererseits die Menge der jetzt verbrauchten Defaults. Das beschreibt den Hauptfall (a) der Definition. (Im Hauptfall gibt es eine Ausstiegsklausel, die greift wenn die bis zu einem Punkt n erweiterte Prämissenmenge mit den bis zu n summierten und den neuen Bedingungen inkonsistent ist. In diesem Fall endet die Konstruktion ohne eine Extension zu erzeugen.) Teil (b) ist die Abschlußklausel der Induktion: Sie behandelt den Fall, daß es in einem Schritt keine neuen anwendbaren Defaults mehr gibt. In diesem Fall bleibt es bei dem im vorangegangenen Schritt erreichten Ergebnis und die Induktion endet hier.

Um diese Idee umzusetzen, beginnen wir also damit, daß wir die Defaults einer Default-Theorie (D, X) nach dem Muster der natürlichen Zahlen anordnen. In einer (strikt) *wohlgeordneten* Menge von Defaults, $(D, \prec)$, ist die Relation $\prec \subseteq D^2$ irreflexiv und transitiv, sowie vollständig,

$$d_1 \prec d_2 \text{ oder } d_2 \prec d_1,$$

und fundiert,

$$\forall (\emptyset \neq) D' \subseteq D : D' \text{ enthält ein } \prec -\text{kleinstes Element.}$$

Eine Default-Theorie mit wohlgeordneten Defaults notieren wir mit '$(D, X, \prec)$'. Ferner wollen wir die Menge aller Bedingungen einer Default-Menge D mit '$Bed(D)$' notieren.

DEFINITION 22. (Makinson) Es sei $(D, X, \prec)$ eine wohlgeordnete Default-Theorie. Wir definieren zwei Folgen von Mengen, X_n und D_n $(0 \leq n < \omega)$.

Basis $n = 0$:

$$X_0 = \mathrm{Cn}(X),$$
$$D_0 = \emptyset.$$

Schritt $n + 1$:

(a) Angenommen, $(*)$ es gibt einen Default (x, Y, z) in D, aber noch nicht in D_n so, daß $x \in X_n$ und $X_n, y \nvdash \bot\, (\forall y \in Y)$ (d.h. der Default ist auf X_n anwendbar). Es sei $d = (A, \mathcal{B}, C)$ der erste Default (unter $\prec$), der diese Bedingungen erfüllt.

(Erweiterung) Falls $X_n, B, C \nvdash \bot$ $(\forall B \in \mathit{Bed}(D_n) \cup \mathcal{B})$, dann

$$X_{n+1} = \mathrm{Cn}(X_n \cup \{C\}),$$
$$D_{n+1} = D_n \cup \{d\}.$$

(Ausstieg) Anderenfalls sind X_{n+1} und D_{n+1} undefiniert, die Konstruktion endet hier und es gibt für $(D, X, \prec)$ keine Erweiterungen.

(b) Angenommen, $(*)$ trifft nicht zu. Dann

$$X_{n+1} = X_n,$$
$$D_{n+1} = D_n.$$

Die Menge $\bigcup\{X_n : n < \omega\}$, falls definiert, ist *die Makinson-Erweiterung* von $(D, X, \prec)$. (Ende Def.)

SATZ 23. *(Makinson) Eine Formelmenge E ist genau dann eine Reiter-Erweiterung von (D, X), wenn es eine Wohlordnung $\prec$ von D gibt so, daß E die Makinson-Erweiterung von $(D, X, \prec)$ ist.*

BEWEIS. Makinson [200, Anhang A]. ■

Beispiele:

1. *Ausstieg.* Es sei $X = \{P\}$ und $d = (P : \neg Q \Rightarrow Q)$ sei der einzige Default. Dann $X_0 = \mathrm{Cn}(P)$ und $D_0 = \emptyset$. Im Schritt für X_1 erfüllt d die Anwendbarkeitsbedingung $(*)$ unter (a), denn $d \notin D_0$ und $P, \neg Q \nvdash \bot$. Aber $P, Q, \neg Q \vdash \bot$. So kommen wir in die Ausstiegsklausel von (a), d.h. X_1 ist nicht definiert und also hat diese Default-Theorie keine Erweiterung. (Wenn wir in diesem Beispiel P durch $\top$ ersetzen und also X leer lassen können, dann entsteht wieder das schon zuvor beschriebene

minimale Beispiel für einen frühen Ausstieg.) — Das Beispiel läßt sich verallgemeinern: An jedem Punkt n der Konstruktion, an dem $A \in X_n$ und X_n konsistent mit $\neg B$ ist, erzwingt ein Default der Form $A : \neg B \Rightarrow B$ den Ausstieg, da B nicht den Test $X_n, \neg B, B \nvdash \bot$ besteht.

2. *Früher Abschluß.* Es sei $X = \{P\}$ und $d = (P : \neg P \Rightarrow Q)$ der einzige Default. X_0 und D_0 sind wie im ersten Beispiel. Für X_1 stellen wir fest, daß zwar $d \notin D_0$, jedoch $P, \neg P \vdash \bot$. So werden wir auf die Abschlußklausel (b) verwiesen: Die gesuchte Erweiterung ist $X_1 = X_0 = \mathrm{Cn}(P)$. (Mit $X = \emptyset$ und dem einzigen Default $\top : \bot \Rightarrow Q$ haben wir ein minimales Beispiel für den frühen Abschluß, mit der einzigen Erweiterung $\mathrm{Cn}(\emptyset)$.)
Defaults der Form $A : \neg A \Rightarrow B$ (oder allgemeiner: $A : \mathcal{B} \Rightarrow B$ mit $\neg A \in \mathcal{B}$) erzwingen offensichtlich an jedem Punkt der Konstruktion den Abschluß, da A und $\neg A$ inkonsistent sind und wir so in die Abschlußklausel geraten.

3. *Hauptfall.* Es sei $X = \{P\}$ und $D = \{P : Q \Rightarrow Q,\ P : \neg Q \Rightarrow \neg Q\}$; d_1 sei der erste, d_2 der zweite Default in D. Nun gibt es zwei mögliche Ordnungen, $d_1 \prec d_2$ und $d_2 \prec d_1$: wir betrachten die erste. Für X_1 stellen wir fest, daß $d_1 \notin D_0$, $P \in X_0$ und $P, Q \nvdash \bot$; d_1 erfüllt also die Bedingung $(*)$ und entgeht ebenfalls der Ausstiegsklausel. So ist $X_1 = \mathrm{Cn}(P, Q)$ und $D_1 = \{d_1\}$. Wir kommen zu X_2: Der einzige noch nicht verbrauchte Default d_2 erfüllt nicht die Bedingung $(*)$, da $X_1, \neg Q \vdash \bot$. Also schließen wir nach (b) ab: $X_2 = X_1 = \mathrm{Cn}(P, Q)$ ist die gesuchte Erweiterung unter der Ordnung $d_1 \prec d_2$. Unter der zweiten der möglichen Anordnungen, erhalten wir dagegen die Erweiterung $\mathrm{Cn}(P, \neg Q)$.

4. *Inkonsistente Prämissen.* Angenommen $X \vdash \bot$. Im ersten Schritt haben wir $X_0 = \mathrm{Cn}(X) = \mathrm{FML}$. Im nächsten Schritt stellen wir fest, daß die Bedingung $(*)$ nicht erfüllt ist und schließen daher nach (b) die Konstruktion der Erweiterung ab. Hier kommt es gar nicht darauf an, welche Defaults zu Verfügung stehen. Es gilt daher ganz allgemein:

Merke: Jede Default-Theorie (D, X) *mit inkonsistenter Prämissenmenge* X *hat genau eine Erweiterung, nämlich* $\mathrm{Cn}(X) = FML$.

5. *Konsistente Prämissen.* Die Menge X sei konsistent und D sei entweder leer oder enthalte nur Defaults, die auf $\mathrm{Cn}(X)$ nicht anwendbar sind. Dann ist offenbar $\mathrm{Cn}(X)$ die einzige Erweiterung von (D, X). Nun nehme man an, daß D anwendbare Defaults enthält. Dann entstehen in jedem Schritt nur konsistente Mengen. Es gilt also allgemein;

Merke: Jede Default-Theorie (D, X) *mit konsistenter Prämissenmenge* X *hat nur konsistente Erweiterungen.*

Die Rolle der Bedingungen $\mathcal{B}$ in einem Default $(A, \mathcal{B}, C)$ ist es, die Anwendung des Defaults durch eine Konsistenzprüfung zu regulieren. Wenn die Prämisse A gegeben ist, das Resultat aber mit den Bedingungen $\mathcal{B}$ unverträglich ist, dann sollten wir den Default $(A, \mathcal{B}, C)$ *nicht* anwenden. Normalerweise erwarten wir, daß unter den Bedingungen $\mathcal{B}$ sich auch die Konklusion C befinden sollte: Wir sollten aus A nicht auf C schließen, wenn das Resultat mit A selbst unverträglich ist. Normalerweise erwarten wir ebenso, daß $\neg C$ keine Bedingung für den Schluß auf C ist, denn das Resultat des Schlusses auf C kann nicht mit $\neg C$ kompatibel sein. (Solche Defaults führen zum Ausstieg, wie das erste Beispiel oben zeigt.)

Was wir normalerweise von Defaults erwarten, das erfüllen die sogenannten *normalen Defaults* von der Form (A, C, C), bzw. allgemeiner $(A, \mathcal{B}, C)$ mit $\mathcal{B} \equiv \{C\}$. Normale Defaults sind, falls überhaupt anwendbar, *konsistent* anwendbar – das macht sie "normal". Wenn wir die Definition 22 für den Fall normaler Defaults reformulieren, dann entfällt so die Ausstiegsklausel und Extensionen sind trivialerweise garantiert. (Im schlimmsten Fall gibt es keine auf X anwendbaren Defaults in D. Dann ist $\mathrm{Cn}(X)$ die einzige Extension von (D, X).) Für normale Defaults ist es auch nicht mehr nötig, über bereits angewandte Defaults Buch zu führen. So erhalten wir aus der Definition 22 die folgende.

Definition 24. Es sei $(D, X, \prec)$ eine wohlgeordnete, *normale* Default-Theorie. Wir definieren eine Folge von Mengen, X_n $(0 \leq i < \omega)$.

Basis $n = 0$:

$$X_0 = \mathrm{Cn}(X).$$

Schritt $n + 1$:

(a) Angenommen, $(*)$ es gibt einen Default $x \dot{\Rightarrow} y$ in D so, daß $x \in A_n$ und $A_n, y \not\vdash \bot$. Es sei $d = A \dot{\Rightarrow} C$ der erste Default (unter $\prec$), der diese Bedingungen erfüllt. Dann

$$X_{n+1} = \mathrm{Cn}(X_n \cup \{C\}).$$

(b) Angenommen, $(*)$ trifft nicht zu. Dann

$$X_{n+1} = X_n.$$

Die Menge $\bigcup\{X_n : n < \omega\}$ ist *die Makinson-Erweiterung* von $(D, X, \prec)$.

Satz 25. *Eine Formelmenge E ist genau dann eine Reiter-Erweiterung einer* normalen *Default-Theorie (D, X), wenn es eine Wohlordnung $\prec$ von D gibt so, daß E die Makinson-Erweiterung von $(D, X, \prec)$ ist.*

Wir schließen diesen Abschnitt mit einer Beobachtung, die uns einen schnellen Zugang zu allgemeinen Eigenschaften des Schließens mit Defaults eröffnet.

- $Anw(D, X)$ bezeichne die Menge der auf X anwendbaren Defaults in D (cf. p. 460).
- $Erw(D, X)$ bezeichne die Menge der Erweiterungen von (D, X) (nach einer der drei äquivalenten Definitionen).

Jede Erweiterung E einer Menge X unter Defaults D ist der Abschluß von X zusammen mit den Konsequenzen aller auf E anwendbaren Defaults.

Beobachtung 26. (Reiter) *Wenn* $E \in Erw(D, X)$*, dann*

$$E = \mathrm{Cn}(X \cup \{C : (A, \mathcal{B}, C) \in Anw(D, E)\}).$$

Beweis. Nach Satz 25 ist E die Makinson-Erweiterung von $(D, X, \prec)$ für eine bestimmte Ordnung $\prec$. Es genügt zu zeigen, daß für jeden Schritt X_n in der Konstruktion von E,

$$(*) \qquad X_n = \mathrm{Cn}\{X \cup \{C : (A, \mathcal{B}, C) \in Anw(D_n, X_n)\}\}.$$

Da die Konstruktion von $E = \bigcup X_n$ $(0 \leq n \leq \omega)$ kumulativ ist, hat E dann selbst die Eigenschaft $(*)$.

Die Basis X_0 ist trivial. Wir nehmen $(*)$ als Induktionsannahme (IA) und betrachten X_{n+1}.

Der Fall (b) der Definition, $X_{n+1} = X_n$, wird durch die IA gedeckt. Im verbleibenden Fall (a) können wir nach der Annahme der Beobachtung $(Erw(D, X) \neq \emptyset)$ die Ausstiegsklausel unberücksichtigt lassen. So ist nach der Definition $X_{n+1} = \mathrm{Cn}(X_n \cup \{C\})$, wobei C die Konklusion ist eines auf $X_n \subseteq X_{n+1}$ anwendbaren Defaults $(A, \mathcal{B}, C) \in D_{n+1}$; d.h. $(A, \mathcal{B}, C) \in Anw(D_{n+1}, X_{n+1})$, wie gewünscht. ∎

8. Folgern aus Erweiterungen

Wir wissen nun, wie wir eine Formelmenge X so erweitern können, daß wir maximal konsistenten Gebrauch von einer gegebenen Menge D von Defaults machen. Was *folgt* nun aus einer Menge von Prämissen unter Rückgriff auf eine Menge von Defaults?

Es sei $Erw(D, X)$ die Menge der Erweiterungen einer Default-Theorie (D, X). Wie zuvor haben wir im Prinzip drei Optionen, aus $Erw(D, X)$ zu folgern: Wir können alle, nur die beste oder eine Auswahl von Erweiterungen in Betracht ziehen. In der bisher verwendeten Terminologie: Wir können aus $Erw(D, X)$ nach vollem Schnitt, einer Maximalwahl oder einem Teilschnitt schließen. (Die Möglichkeit Erweiterungen zu vereinigen, scheidet auch hier aus, da alle Erweiterungen einer Default-Theorie untereinander inkonsistent sind.)

Definition 27. Eine Formel A ist eine *Default-Konklusion* aus Prämissen X unter Rückgriff auf Defaults D bzw. A ist eine *Default-Folgerung* aus der Default-Theorie (D, X) (Notation: $X \mathrel{|\!\sim}_D A$, auch $A \in \mathrm{C}_D(X)$) gdw ... (es folgen die drei Optionen:)

1. ... A in *allen* Erweiterungen von (D, X) ist; oder
2. ... A in *einer ausgewählten* Erweiterung von (D, X) ist; oder
3. ... wenn A in jeder einer *ausgewählten Teilmenge* von Erweiterungen von (D, X) ist.

In der Literatur hat sich für Folgerungen des Typs 1, d.h. nach dem vollen Schnitt, der Begriff einer *skeptischen* Default-Folgerung etabliert. Folgerungen des Typs 2 werden dagegen *mutig* (*brave*) oder auch "gutgläubig" ("credulous") genannt. Die Konklusionen skeptischen Folgerns aus Defaults sind derart, daß sie aus den Prämissen folgen, gleichgültig welche Extension wir betrachten, d.h. gleichgültig in welcher Reihenfolge die Defaults ausgelöst werden. In anderen Worten, wir können es dem Skeptiker überlassen, eine Reihenfolge zu wählen. Skeptisches und mutiges Schließen sind natürlich Grenzfälle des Typs 3 nach einem Teilschnitt. Dieser dritte Typ wird so abstrakt als Folgerungsart kaum thematisiert. Stattdessen werden gewöhnlich bestimmte Arten, die Auswahl zu treffen, betrachtet und die so definierten Folgerungen untersucht. Wir werden unter dem Stichwort "Priorisierung" kurz darauf zurückkommen.

Unter einer *Default-Logik nach Reiter* (DL) im engeren Sinne verstehen wir die skeptische Folgerungsrelation $\mathrel{|\!\sim}_D$ für eine bestimmte Menge D von Defaults. Im weiteren Sinne sei damit die Theorie solcher Folgerungsrelationen unter Zugrundelegung des Reiter'schen Begriffs der Erweiterung einer Default-Theorie gemeint. Im folgenden werden wir von DL meist in diesem weiteren Sinne sprechen.

Defaults $(A, \mathcal{B}, C)$ können sehr verschiedene Formen annehmen. Das macht DL zu einer besonders flexiblen Theorie. (Wir schreiben abkürzend wieder (A, B, C) für $(A, \{B\}, C)$.)

(a) Normale Defaults (A, C, C) sind genau dann anwendbar, wenn das Resultat der Anwendung konsistent ist. Weitere Bedingungen müssen nicht erfüllt sein. Derselbe Effekt wird durch Defaults der Form $(A, \top, C)$ erzielt, da $(A, \top, C)$ und (A, C, C) unter derselben Bedingung konsistent anwendbar sind.

(b) Ein Default $(\top, B, C)$ ist auf beliebige Prämissen anwendbar, solange das Resultat nur mit den Bedingungen konsistent ist. Ein solcher Default ist wie eine Annahme, die unter der Bedingung der Konsistenz mit den Bedingungen verwendet werden kann.

(c) Ein Default $(A, \emptyset, C)$ stellt keine Bedingungen für die Anwendung, außer der, daß das Antezedens A durch die Prämissen gegeben ist; insbesondere findet hier keine Konsistenzprüfung statt. Solche Defaults emulieren also das unvorsichtige (monotone) Folgern unter materialen Regeln nach Definition 14.

(d) Defaults der Form $(\top, C, C)$ (äquivalent $(\top, \top, C)$) kombinieren die Fälle (a) und (b). In der Literatur werden sie *supernormale* Defaults genannt. Diese repräsentieren Annahmen, die zu beliebigen Prämissen hinzugenommen werden können, solange nur die Konsistenz des Resultats gewährleistet ist. Das Schließen aus solchen Defaults ist also wie das Schließen aus Default-Annahmen nach Definition 4.

(e) Schließlich kombinieren Defaults der Form $(\top, \emptyset, C)$ die Fälle (b) und (c). Sie erlauben die Annahme C zu beliebigen Prämissen ohne Rücksicht auf Konsistenz hinzuzunehmen. Das entspricht dem Schließen aus materialen Annahmen nach der Definition auf S. 414.

Um die Vielseitigkeit des Formalismus zu illustrieren, wollen wir kurz supernormale Defaults der Form $\top \dot{\Rightarrow} \neg P$, mit P atomar, betrachten. Diese besagen soviel wie: Nimm an, daß $\neg P$, falls das konsistent mit den gegebenen Prämissen ist. Das ist unschwer als eine Formulierung der Annahme einer Geschlossenen Welt (CWA) zu erkennen (vgl. pp. 436ff.). (Zur Erinnerung: $cwa(X) = \{\neg P : X \nvdash X\} \cup X\}$.) Deshalb können wir von

$$D_{cwa} = \{\top \dot{\Rightarrow} \neg P : \ P \in \mathrm{ATM}\}$$

als der Menge der "CW-Defaults" sprechen. Wenn wir D_{cwa} auf eine Menge X anwenden, dann generieren wir genau die Menge X zusammen mit allen Literalen $\neg P$, die mit X konsistent sind. Mit anderen Worten, die Default-Theorie (D_{cwa}, X) hat genau eine Erweiterung, nämlich $\mathrm{Cn}(cwa(X))$. So läßt sich im Rahmen der Default-Logik auch das Schließen nach der CWA darstellen.

Wir haben beobachtet (p. 463), daß die Beschränkung auf normale Defaults die Theorie erheblich vereinfacht und Erweiterungen in jedem Fall garantiert. Wenn wir nun die verschiedenen Formen unter (a–e) betrachten, dann können wir zunächst feststellen, daß der Fall (c) vergleichsweise uninteressant ist und (d) eine Kombination der Grundformen (a–b) ist. Die Form (b) subsumiert normale Defaults der Form $(\top, C, C)$. Was verlieren wir an Ausdrucksstärke, wenn wir uns generell auf die einfacheren normalen Defaults beschränken? Um diese Frage zu beantworten, betrachten wir das klassische Pinguin-Szenario.

Beispiel: Normalerweise fliegen Vögel, es sei denn, es handelt sich um Pinguine. Pinguine fliegen normalerweise nicht. Was dürfen wir schließen aus den Annahmen, daß Pinguine Vögel sind und Robert ein Pinguin ist? Es gibt hier verschiedene, nicht äquivalente Möglichkeiten, diese Situation mit Defaults darzustellen.

1. *Normale Defaults*:

   $Q \dot{\Rightarrow} R$ ("Vögel fliegen normalerweise").

   $P \dot{\Rightarrow} \neg R$ ("Pinguine fliegen normalerweise nicht").

Es sei nun $X = \{P \to Q, P\}$. Dann gibt es zwei Erweiterungen von X unter diesen Defaults, $E_1 = \{P, P \to Q, R\}$ und $E_2 = \{P, P \to Q, \neg R\}$. Im Sinne des skeptischen Schließens aus Defaults folgt also nicht, daß $\neg R$ ("Robert fliegt nicht"). (Da wir hier vereinfachend auf die prädikatenlogische Darstellung verzichten, stehen die Paraphrasen in Anführungszeichen. Der geneigte Leser weiß, wie es gemeint ist.)

2. *Default mit Ausnahmen*:

   $Q : R \wedge \neg P \Rightarrow R$ ("Vögel, außer Pinguine, fliegen normalerweise")

   $P : \neg R \Rightarrow \neg R$ ("Pinguine fliegen normalerweise nicht")

X sei wie zuvor. Jetzt kann jedoch der erste Default nicht ausgelöst werden, weil die Ausnahme P vorliegt. Deshalb gibt es nur eine Erweiterung von X, nämlich E_1. Wir dürfen also schließen, daß $\neg r$ ("Robert fliegt nicht").

Die Formalisierung 1 im Beispiel benutzt nur normale Defaults; sie kann aber nicht das gewünschte Ergebnis erzeugen. Dagegen machen wir in der Formalisierung 2 von einem nicht-normalen Default Gebrauch, in dem wir eine bestimmte Ausnahme von der Regel, daß Vögel fliegen, benennen. Mit dieser Formalisierung bekommen wir das richtige Resultat. Natürlich gibt es unendlich viele mögliche Ausnahmen von der Regel, daß Vögel fliegen. Im allgemeinen können wir diese Möglichkeiten nicht alle in den Bedingungen eines Defaults benennen. Aber in kontextuell oder explizit definierten Anwendungssituationen kann es sinnvoll sein, für eine Regel nur eine begrenzte Anzahl von Ausnahmen vorzusehen. In solchen Fällen möchten wir, wie im Beispiel oben, Defaults mit expliziten Ausnahmen, d.h. Defaults der Form $A : B \wedge C \Rightarrow C$ formulieren. Solche Defaults heißen *seminormal*: Sie fordern, wie normale Defaults, die Konsistenz mit der Konklusion C, jedoch auch weitere zu erfüllende Bedingungen B. (Allgemeiner: $A : \mathcal{B} \Rightarrow C$ ist seminormal, falls $\mathcal{B} \vdash C$.) Seminormale Defaults subsumieren offensichtlich normale Defaults, können aber darüber hinaus Regeln ausdrücken, die mit normalen Defaults nicht ausdrückbar sind. Auch $A : \neg C \wedge C \Rightarrow C$ ist ein seminormaler Default. Von solchen Defaults wissen wir, daß sie keine Erweiterungen erzeugen können. Seminormale Defaults führen daher zu keiner wesentlichen Vereinfachung der Theorie, so wie das normale Defaults tun.

Wir haben soeben gesehen, daß die Logik seminormaler Defaults strikt ausdrucksstärker als die Logik normaler Defaults ist. Das war kein überraschendes Resultat. Wie verhält sich nun die Ausdrucksstärke der Logik seminormaler Defaults zu der beliebiger Defaults? Wonach genau fragen wir hier? Uns interessiert das Schließen aus Annahmen X unter Rückgriff auf Defaults D. Wir möchten jetzt wissen, ob eine Konklusion aus X unter Rückgriff auf Defaults D *beliebiger* Form auch unter Rückgriff auf Defaults D' einer syntaktisch eingeschränkten Teilklasse von Defaults, hier der *seminormalen* Defaults, erreicht werden kann. Gegeben die Definition des Schließens aus Defaults, genügt es offenbar die folgende Frage zu beantworten:

> Gibt es eine Übersetzung f von beliebigen in seminormale Default-Theorien so, daß $Erw(D, X) = Erw(f(D, X))$?

Die Antwort auf die Frage wird auch von den Bedingungen abhängen, die wir an eine Übersetzung stellen. Aber selbst unter starken Bedingungen, fällt die Antwort positiv aus (Janhunen [153, Thm 2]). Default-Logik kann im wesentlichen auf die Logik seminormaler Defaults reduziert werden.

Die Default-Logik nach Reiter ist also eine sehr ausdrucksstarke Theorie, selbst wenn wir uns auf seminormale Defaults beschränken. Defaults können sowohl Regeln als auch Annahmen ausdrücken, beides jeweils mit oder ohne Bedingungen. All diese Arten von Defaults können in *einer* Default-Menge kombiniert werden. Aber nicht nur durch ihre Ausdrucksstärke, läßt sich Default-Logik an verschiedene Anwendungsbedürfnisse flexibel anpassen. Auch lassen sich die Kerndefinitionen (Erweiterung, Schließen aus Erweiterungen) relativ umstandslos variieren, um gewünschte Eigenschaften zu garantieren, wie wir im nächsten Abschnitt sehen werden.

Nach unserer Beobachtung 26 sind alle Erweiterungen E einer Default-Theorie (D, A) in der Form $\mathrm{Cn}(A \cup \mathrm{kon}(D'))$ darstellbar, wobei $\mathrm{kon}(D')$ die Menge der Konklusionen aus einer Teilmenge $D' \subseteq D$ von Defaults – nämlich den in E angewandten – ist. Das bedeutet im Falle des skeptischen Schließens, daß $\mathrm{C}_D(A)$ der Schnitt solcher Mengen $\mathrm{Cn}(A \cup \mathrm{kon}(D'))$ ist. Dieser Beobachtung lassen sich schon einige der wichtigsten strukturellen Eigenschaften der Operation C_D (bzw. der Relation $\mid\!\sim_D$) entnehmen. So läßt sich z.B. schnell nachprüfen, daß $\mid\!\sim_D$ rechts deduktiv abgeschwächt werden kann und auch die Schnitteigenschaft (kumulative Transitivität) hat:

$$\frac{X \mid\!\sim_D A \quad A \vdash B}{X \mid\!\sim_D B} \qquad \frac{X \mid\!\sim_D A \quad X, A \mid\!\sim_D B}{X \mid\!\sim_D C}.$$

Jedoch ist $\mid\!\sim_D$ weder allgemein kompakt noch erfüllt die Relation einige andere wünschenswerte Bedingungen. Zu solchen Bedingungen gehören sicher die Disjunktivität in den Prämissen und die Vorsichtige Monotonie,

$$\frac{X, A \mid\!\sim_D C \quad X, B \mid\!\sim_D C}{X, A \vee B \mid\!\sim_D C} \qquad \frac{X \mid\!\sim_D A \quad X \mid\!\sim_D B}{X, B \mid\!\sim_D A}.$$

Beispiele: 1. Für Disjunktivität können wir das Beispiel von p. 452 adaptieren. Es sei $D = \{P : \top \Rightarrow R,\, Q : \top \Rightarrow R\}$. Dann $\mathrm{C}_D(P) = \mathrm{Cn}(P, R)$, $\mathrm{C}_D(Q) = \mathrm{Cn}(Q, R)$ und $\mathrm{C}_D(P \vee Q) = \mathrm{Cn}(P \vee Q)$. Während also $P \mid\!\sim_D R$ und $Q \mid\!\sim_D R$, gilt nicht $P \vee Q \mid\!\sim_D R$.

2. Für Vorsichtige Monotonie dient folgendes, kleines Gegenbeispiel (Makinson). Es sei $D = \{P : \top \Rightarrow \neg Q,\, Q : \top \Rightarrow P,\, \top : \top \Rightarrow Q\}$. Dann ist $Erw(D, \{\top\}) = \{\mathrm{Cn}(P, Q)\}$ und $Erw(D, \{P\}) = \{\mathrm{Cn}(P, Q), \mathrm{Cn}(P, \neg Q)\}$. So haben wir $\top, P \not\mid\!\sim_D Q$ obwohl $\top \mid\!\sim_D P$ und $\top \mid\!\sim_D Q$.

9. Default-Logik: Probleme und Varianten

In diesem Abschnitt sollen einige Eigenschaften der DL nach Reiter besprochen werden, die möglicherweise als Unzulänglichkeiten interpretiert werden können und so zur Entwicklung von Varianten der DL Anlaß geben. Am Ende dieses Abschnitts steht jedoch ein Resultat, welches einmal mehr die Flexibilität der DL unter Beweis stellt: In einem bestimmten Sinne sind die scheinbar konkurrierenden Varianten in der klassischen DL enthalten.

Wenn wir DL als einen spezifischen Vorschlag für die Formalisierung vorsichtig materialen Schließens verstehen, dann scheint die Tatsache, daß DL weder unter Prämissendisjunktion noch unter Vorsichtiger Monotonie abgeschlossen ist, klar gegen diesen Vorschlag zu sprechen. Aber so einfach ist die Sache nicht. Das Schließen aus Defaults kann nur so gut sein wie die Defaults, aus denen geschlossen wird. Wenn eine Menge D von Defaults bestimmte Qualitätsmerkmale nicht erfüllt, dann können problematische Eigenschaften des Folgerns aus solchen problematischen Defaults nicht überraschen. Mit dieser Überlegung empfiehlt sich eine argumentative Option, die wir ähnlich auch im Hinblick auf die deduktive Folgerungsrelation $\vdash$ verfolgen können. Danach reflektiert beispielsweise die "Pathologie" der *ex falso quodlibet*-Eigenschaft $A, \neg A \vdash B$ nicht einen Defekt in der Definition der Folgerungsrelation, sondern einen in den Prämissen.

Im zuletzt vorgeführten Gegenbeipiel 1 zur Prämissendisjunktion ist die Default-Menge in einem offensichtlichen Sinne unvollständig: Zwar haben wir $P \Rightarrow R$ (hier abkürzend für $P : \top \Rightarrow R$) und $Q \Rightarrow R$, jedoch fehlt der Default $P \vee Q \Rightarrow R$. D.h. die Default-Menge selbst ist nicht unter Oder-Einführung in den Prämissen der $\Rightarrow$-Relation abgeschlossen. Dieser Umstand spiegelt sich erwartbar in der Relation $\mid\!\sim_D$ wieder. Das Beispiel zeigt auf besonders einfache Weise, daß wir gewünschte Eigenschaften manchmal garantieren können, in dem wir die Klasse zulässiger Default-Mengen unter entsprechende Bedingungen stellen. In diesem Fall wäre das die Bedingung, daß Defaultmengen unter Prämissendisjunktion für $\Rightarrow$ abgeschlossen sind. Statt den Anwendungsbereich der Definition einer Erweiterung einzuschränken, können wir aber auch in die Definition selbst eingreifen, indem wir fordern, daß für jede Disjunktion in einer Prämissenmenge X, alle Erweiterungen von X entweder das eine oder das andere Disjunkt enthalten (sogenannte *Disjunctive DL*; siehe [105]).

Ähnlich verhält es sich mit dem Gegenbeispiel 2 (siehe oben) zur Vorsichtigen Monotonie. Hier haben wir eine Default-Menge D bestehend aus $\Rightarrow q, q \Rightarrow p$ und $p \Rightarrow \neg q$ (wieder in der vereinfachten, die leere Bedingung $\top$ fortlassende Notation). Nun hat D die merkwürdige Eigenschaft, daß es gar keine Menge von Prämissen geben kann, auf welche alle Defaults in D konsistent angewandt werden können: Unabhängig von jeder Prämissenmenge ist D schon selbst derart, daß die Defaults sich gegenseitig blockieren. Damit ist eine unabhängig motivierte Bedingung für zulässige Default-Mengen angedeutet, die Gegenbeispiele zur Vorsichtigen Monotonie ausschließt. Auch hier können wir alternativ die Definition einer Erweiterung so ändern, daß Vorsichtige Monotonie allgemein erfüllt ist. Die *Cumulative DL* nach Brewka [39] ist das Resultat eines solchen Eingriffs in die Definition von Reiter.

Wie sollen wir mit dem Umstand umgehen, daß für einige Default-Theorien Erweiterungen nicht definiert sind? Die Beispiele für solche Fälle (siehe p. 461) zeigen Defaults wie $A : \neg B \Rightarrow B$ am Werke, Defaults also, die schon für sich genommen nicht konsistent anwendbar sind. Wir können, wie bereits im Falle der Prämissendisjunktion und der Vorsichtigen Monotonie vorgeschlagen wurde, solche pathologischen Defaults von der Bildung zulässiger Default-Theorien ausschließen. Alternativ können wir in die Definition einer Extension so eingreifen, daß Extensionen immer garantiert sind. Für diese Alternative spricht aus theoretischer Perspektive, daß sie allgemeiner, aus praktischer, daß sie fehlertoleranter ist. Versionen der Default-

Logik, in der Erweiterungen immer definiert sind, werden in [189] (*Justified DL*) und [61, 256] (*Constrained DL*) vorgeschlagen.

In der DL-Definition einer Erweiterung wird für die Anwendbarkeit eines Defaults $A : \mathcal{B} \Rightarrow C$ gefordert, daß jede einzelne Bedingung $B \in \mathcal{B}$ mit der Erweiterung $E \subseteq$ FML konsistent sein soll: $E, B \not\vdash \bot$. Das schließt nicht aus, daß E mit der Gesamtheit der Bedingungen inkonsistent ist: $E, \mathcal{B} \vdash \bot$. ($E$ sei beispielsweise unvollständig bezüglich B, d.h. $E \not\vdash B$ und $E \not\vdash \neg B$, und $\mathcal{B}$ enthalte B und $\neg B$.) Im allgemeinen ist es daher nicht so, daß ein Default der Form $(A, \{B_1, B_2\}, C)$ denselben Effekt hat wie $(A, \{B_1 \wedge B_2\}, C)$. Der zweite Default ist schwächer als der erste, da jener weniger Anwendungsfälle als dieser hat. DL unterscheidet also zwischen logisch äquivalenten Bedingungsmengen. Wir können auch dies als ein Zeichen der Flexibilität der DL sehen. Würden wir statt der Konsistenzprüfung jeder einzelnen Bedingung die summarische Prüfung der Bedingungsmenge fordern, so würde diese Flexibilität verloren gehen. In solchen Fällen, in denen wir die summarische Prüfung einer Bedingungsmenge haben wollen, können wir sie (im endlichen Fall) im Rahmen der Reiter-Definition simulieren, indem wir den konjunktiven Abschluß der Bedingungen prüfen.

Wie nun schon angedeutet, stellt sich heraus, daß die gewünschten Eigenschaften, welche bestimmte Varianten der DL auszeichnen, im Rahmen der DL selbst bei geeigneter Wahl von Annahmen und Defaults erzielt werden können. Alle hier betrachteten Varianten verändern die Definition der Erweiterung einer Default-Theorie.

> Es sei $Erw^V(D, X)$ die Menge der Eweiterungen von (D, X) im Sinne der neuen Definition. Dann läßt sich (wie Delgrande und Schaub [63] zeigen) eine Übersetzung (D^*, X^*) von (D, X) angeben, so daß $Erw^V(D, X) = Erw(D^*, X^*)$, wobei Erw die Erweiterungen im Sinne der Reiter-Definition bestimmt. Ein solches Resultat läßt sich für *disjunctive, cumulative, justified, constrained*, sowie für die meisten anderen bekannten Varianten der DL erzielen.

Diese wichtige Reihe von Resultaten legt nahe, DL nicht als einen spezifischen Vorschlag für die Definition einer Folgerungsrelation zu verstehen, zu dem verschiedene Varianten in Konkurrenz treten können. Vielmehr handelt es sich bei der DL um einen sehr allgemeinen theoretischen Rahmen, innerhalb dessen spezifische Vorschläge formuliert und verglichen werden können – etwa so, die Kripke-Semantik eine Blaupause für die Modelltheorien spezifischer Modallogiken darstellt. Daß die allgemeine Theorie recht schwache Eigenschaften hat, kann aus dieser Sicht nicht überraschen, sondern belegt nur die wünschenswerte Allgemeinheit der DL.

Priorisierte Defaults. Wir haben gesehen, daß DL eine sehr ausdrucksstarke Theorie ist unter die viele bekannte Ansätze subsumiert werden können. (Einige davon, wie Moores Autoepistemische Logik [215], haben wir hier gar nicht besprochen.) Aber stellt DL alle Resourcen zur Verfügung, die für das Modellieren vorsichtig materialen Schließens relevant sein können? Ein Parameter, der für solches Schließen eine wichtige Rolle spielen kann, ist in der DL nach Reiter nicht präsent: die relative Gewichtung von Defaults. Interessante Default-Theorien sind derart, daß in einer Erweiterung nicht alle Defaults ausgelöst werden können. Solche Default-Theorien führen zu einer Vielzahl von Erweiterungen. Offenbar ist die Frage, ob eine Erweiterung im Einklang oder entgegen einer Gewichtung von Defaults zustande gekommen ist, von höchster Wichtigkeit für die Frage, welche Erweiterungen für das Schließen überhaupt berücksichtigt werden sollen.

Wenn wir Erweiterungen bilden, dann sollten wir, falls möglich, zuerst die vorrangigen Defaults auslösen. Erweiterungen, welche die Vorrangsordnung unter den Defaults nicht respektieren, sollten wir ignorieren. Wenn wir die Menge der Defaults durch eine Ordnung strukturieren, dann kommen wir so zu einer Folgerungsrelation, die nach dem Muster des Teilschnitts bestimmt ist; siehe Definition 27.3. (Einen direkteren Weg zu Teilschnitten, den wir hier nur erwähnen ohne ihn weiter zu behandeln, eröffnet eine Ordnung der Erweiterungen einer Default-Theorie – statt der Defaults selbst –, so daß sich maximale Elemente in dieser Ordnung auszeichnen lassen.)

Bevor wir die Betrachtung gewichteter Defaults fortsetzen, sollten wir nicht unbeachtet lassen, daß die klassische DL über genügend Resourcen verfügt, den Effekt gewichteter Defaults in Einzelfällen hervorzubringen. Ein einfaches Beispiel soll hier zur Illustration genügen.

> *Beispiel*: Es sei $X = \{A, B\}$ und $D = \{(A, C, C), (B, \neg C, \neg C)\}$. Angenommen $d_1 = (A, C, C)$ hat den Vorrang vor $d_2 = (B, \neg C, \neg C)$. Dann wird d_1, nicht jedoch d_2 ausgelöst. So wäre $\{A, B, C\}$ die einzige Erweiterung von (D, X).
> Derselbe Effekt läßt sich erzielen, indem wir den Vorrang von d_1 vor d_2 in eine zusätzliche Bedingung für d_2 kodieren. Diese Bedingung fordert d_2 nur dann auszulösen, wenn d_1 nicht ausgelöst werden kann. D.h. wir übersetzen d_2 zu $d_2^* = (B, \neg A \wedge \neg C, \neg C)$. Nun hat $(D^* = \{d_1, d_2^*\}, X)$ dieselbe Erweiterungsmenge wie (D, X).

In komplexeren Szenarien erfordert es offensichtlich erheblichen Aufwand, eine Default-Ordnung gleichwertig in die Bedingungen einer Default-Menge zu kodieren. Auch wenn es im Prinzip möglich sein sollte priorisierte Default-Theorien (s. Definition unten) in klassische zu übersetzen – ein solches allgemeines Resultat liegt jedoch nicht vor –, so ist die Betrachtung gewichteter Defaults einfacher und intuitiv völlig plausibel.

Wie sollen wir uns die Bestimmung einer Vorrangordnung der Defaults vorstellen? Der einfachere Ansatz geht von einer fixen Ordnung der Defaults aus, die in allen Folgerungen aus Prämissen unter Rückgriff auf diese Defaults unverändert bleibt. Der komplexere Ansatz erlaubt es, die Default-Ordnung selbst durch anfechtbares Schließen auszuhandeln. Wir betrachten zunächst den einfacheren Ansatz.

Den einfacheren Ansatz betrachten wir zunächst unter der weiter vereinfachenden Annahme, daß die Vorrangordnung der Defaults, $(D, \prec)$, eine strikte Wohlordnung ist, d.h. so wie in der Definiton 22 einer Makinson-Erweiterung. Um an die in der Literatur verbreitete Terminologie anzuschließen, wollen wir sagen, daß $(D, \prec)$ eine (unter $\prec$) *wohlpriorisierte* Menge von Defaults ist. Die Aufgabe ist es nun, die Menge der Erweiterungen der Prämissen X unter Rückgriff auf die wohlpriorisierten Defaults $(D, \prec)$, $Erw(D, X, \prec)$ zu bestimmen. Das haben wir in Definition 22 schon getan. Wenn die Theorie $(D, X, \prec)$ überhaupt eine Erweiterung hat, dann hat sie genau eine (welche *per definitionem* unter Konsequenz abgeschlossen ist). Also können wir die Relation der *Folgerung unter Rückgriff auf eine Menge wohlpriorisierter Defaults* so definieren:

$$X \mid\!\sim_{D\prec} A \text{ gdw } A \in Erw(D, X, \prec).$$

Nun verstehen wir unter einer Ordnung des Vorrangs von Defaults eine solche, die in einer gegebenen Situation einem Default den Vorzug, d.h. die *Präferenz* vor einem anderen gibt. Eine Wohlordnung ist ein extremer und selten anzutreffender Fall einer Präferenzordnung. Nicht immer läßt sich von je zwei Optionen sagen, welche die bessere sei. Um (strikte) Präferenz in dem nötigen allgemeinen Sinne zu erfassen, sollten wir nur fordern, daß die Relation irreflexiv (d.h. strikt) und transitiv ist. Es sei $(D, <)$ eine nach Präferenzen in diesem allgemeineren Sinne geordnete, d.h. *priorisierte* Menge von Defaults. Jede Präferenzordnung $<$ läßt sich offenbar, und in der Regel auf verschiedene Weise, zu einer Wohlordnung $\prec$ erweitern. Wenn wir umgekehrt die Familie aller Wohlordnungen betrachten, die eine gegebene Präferenzordnung erweitern, dann ist deren Schnitt genau diese Präferenzordnung: Die spezifische Weise der Erweiterung kürzt sich bei der Betrachtung aller Erweiterungen heraus. Diese einfache Beobachtung erlaubt es, den allgemeineren Fall priorisierter Defaults auf den einfacheren Fall wohlpriorisierter Defaults zurückzuführen. Wenn wir sagen wollen, was aus einer priorisierten Menge D von Defaults folgt, dann genügt es alle Wohlordnungen von D zu betrachten, die mit der gegebenen Priorisierung übereinstimmen, d.h. diese enthalten. Es sei also

$$Erw(D, X, <) = \{Erw(D, X, \prec) : < \subseteq \prec\}.$$

Dann ist die *Folgerung unter priorisierten Defaults* so definiert

$$X \mathrel{|\!\sim}_{D<} A \text{ gdw } A \in \bigcap Erw(D, X, <);$$

oder äquivalent:

$$X \mathrel{|\!\sim}_{D<} A \text{ gdw } \forall \prec \supseteq <: \; X \mathrel{|\!\sim}_{D\prec} A.$$

Das ist im wesentlichen die *Priorisierte DL* (PDL) von Brewka [38, 40, 41], die zugleich Ausgangspunkt für die Dynamisch Priorisierte DL (DPDL) ist, welcher wir uns jetzt zuwenden wollen.

Normalerweise gewinnt Bayern-München ein Bundesliga-Spiel (der Bayern-Default, d_m). Das galt, zumindest eine Zeit lang, auch für Borussia Dortmund (der Borussia-Default, d_b). Was passiert, wenn Borussia gegen Bayern spielt? Bayern-Fans neigen dazu anzunehmen, daß der eine Default, Bayern gewinnt, etwas stärker als der andere Default sei. Wir können diese Annahmen so ausdrücken: $d_b < d_m$. Wenn wir nun den Fall betrachten, daß Bayern gegen Borussia spielt, dann werden Bayern-Fans schließen, daß Borussia nicht gewinnt. Was geschieht, wenn Bayern in schwacher und Borussia in starker Besetzung spielt (B)? Nun, wenn B, dann gilt die Default-Präferenz $d_b < d_m$ normalerweise nicht. Das können wir selbst als einen (normalen) Default darstellen: $(B, \neg(d_b < d_m), \neg(d_b < d_m))$, bzw. $B \dot{\Rightarrow} \neg(d_b < d_m)$. (Unter der Prämisse B ist der Ausgang des Spiels auch für Bayern-Fans offen.) Das Beispiel führt vor, wie Default-Präferenzen selbst Bedingungen und Konklusionen von Defaults sein können. Es ist nicht schwierig nach diesem Muster Beispiele zu finden, in denen auch die Prämissen von Defaults selbst Defaults enthalten können.

Um solche Beispiele im Rahmen der DL zu behandeln, müssen wir zunächst die Objektsprache so erweitern, daß wir Präferenzen zwischen Defaults ausdrücken können. Sodann müssen wir die Erweiterungen von Default-Theorien (D, X) so bestimmen, daß zwei neue Möglichkeiten berücksichtigt werden: Erstens, können unter den Prämissen X Default-Priorisierungen enthalten sein, und, zweitens, können Priorisierungen selbst bloßen Default-Charakter haben, deren Aktivierung im Schließen abhängig ist von den Annahmen und anderen Defaults und deren Priorisierungen. Die richtige Reihenfolge, in der die Defaults zu aktivieren sind, wird so selbst durch Default-Schließen erst ausgehandelt. Das hört sich vertrackt an. Tatsächlich läßt sich das Problem aber wieder recht einfach auf die Betrachtung von Makinson-Erweiterungen reduzieren.

Defaults $(A, \mathcal{B}, C)$ basieren auf den Ausdrücken unserer formalen Sprache, sie sind aber nicht selbst solche Ausdrücke. Wenn wir *in* unserer formalen Sprache Priorisierungen zwischen Defaults ausdrücken möchten, dann müssen wir die Sprache zunächst so erweitern, daß wir auf Defaults Bezug nehmen können. Dazu definieren wir eine Benennungsfunktion und eine Relation $<$:

1. Wenn d ein Default ist, dann sei $\bar{d}$ ein Default-Name.
2. Wenn $\bar{d}_1$ und $\bar{d}_2$ Default-Namen sind, dann sei $\bar{d}_1 < \bar{d}_2$ eine *Formel.*
3. Eine Menge X von Formeln sei genau dann *unter Präferenz abgeschlossen*, wenn für alle $\bar{d}_{i,j,k}$ gilt:
   $\bar{d}_i < \bar{d}_i \notin X$; und
   wenn $\bar{d}_i < \bar{d}_j \in X$ und $\bar{d}_j < \bar{d}_k \in X$, dann $\bar{d}_i < \bar{d}_k \in X$.

Wir wollen die (kleinste) Erweiterung einer aussagenlogischen Sprache nach Bedingungen 1 und 2 einfach eine *erweiterte Sprache* nennen. Man beachte, daß eine Erweiterung nach 1 und 2 beträchtliche Komplexität erzeugen kann. Zum Beispiel ist nach der Definition $d_1 = (\bar{d}_2 < \bar{d}_3, \mathcal{B}, \bar{d}_4 < \bar{d}_5)$ ein Default. Dieser kann selbst benannt werden um so auf ihn in Prämissen, Bedingungen oder weiteren Defaults Bezug zu nehmen.

Definition 28. Eine *dynamisch priorisierte Default-Theorie* ist eine Default-Theorie (D, X) in einer erweiterten Sprache, in der X unter Präferenz abgeschlossen ist.

Die vorläufige Menge der Erweiterungen einer dynamisch priorisierten Default-Theorie ist definiert wie die Menge der Erweiterungen für gewöhnliche Default-Theorien. (Letztere sind übrigens solche dynamisch priorisierten Default-Theorien, welche von den neuen Ressourcen der erweiterten Sprache keinen Gebrauch machen.) Nun wollen wir aber die in den Annahmen und Defaults enthaltene Information über die Priorisierung von Defaults nutzen, um solche Erweiterungen auszuschließen, welche die so kodierte Priorisierung nicht beachten. Solche Erweiterungen, welche die in D und A ausgedrückten Priorisierungsforderungen beachten ("bevorzugte Erweiterungen", s.u.), sind diejenigen, aus denen wir schließen sollten.

Es sei E eine Erweiterung von (D, X). Dann ist E identisch mit einer Makinson-Erweiterung $Erw(D, X, \prec)$. Nun enthält E im typischen Falle Priorisierungsinformation, d.h. Formeln der Form $\bar{d}_i < \bar{d}_j$. Wenn E erzeugt wurde ohne diese Information zu beachten, dann sollten wir diese Erweiterung nicht weiter berücksichtigen. Aber das bedeutet nichts anderes, als daß die Reihenfolge $\prec$ in der Darstellung von E als Makinson-Erweiterung mit den in E gemachten Vorgaben kompatibel sein sollte. Die folgende Definition setzt diese Idee formal um, indem sie zunächst die Relation $\prec$ objektsprachlich repräsentiert und dann Kompatibilität auf Konsistenz reduziert.

Definition 29.
1. Es sei (D, X) eine dynamisch priorisierte Default-Theorie, und $E = Erw(D, X, \prec)$ die Makinson-Erweiterung von (D, X) unter einer Wohlordnung $\prec$. Dann ist $\prec$ *kompatibel* mit $<$ gdw $E \cup \{\bar{d}_i < \bar{d}_j : d_i \prec d_j\}$ konsistent ist.
2. Eine Erweiterung E von (D, X) ist genau dann *bevorzugt*, wenn E die Makinson-Erweiterung von (D, X) unter einer mit $<$ kompatiblen Ordnung $\prec$ ist.
3. Es sei $Erw^*(D, X)$ die Menge der bevorzugten Erweiterungen von (D, X). Dann $X \mid\!\sim_D A$ (*Folgerung unter priorisierten Defaults*) gdw $\forall E \in Erw^*(D, X) : A \in E$.

Damit ist die *Dynamisch Priorisierte DL* (DPDL) von Brewka [40, 41] definiert. Diesem Ansatz folgt im wesentlichen auch Horty [139, 138, 140], mit einem Unterschied: Sowohl in der PDL als auch in der DPDL werden die Erweiterungen einer Default-Theorie letztlich durch Quantifikation über bestimmte Makinson-Erweiterungen wohlgeordneter Default-Theorien bestimmt. In Hortys Theorie übernehmen Erweiterungen, die auf "proper scenarios" basieren, die Rolle der Makinson-Erweiterungen. Auch *proper scenarios* werden durch eine totale Ordnung der Defaults bestimmt (Fundierung ist nicht nötig). Die Ordnung regelt aber nicht einfach die Reihenfolge der Anwendung der Defaults, sondern wird in einer etwas anderen Weise für die Konstruktion von *proper scenarios* und damit von Erweiterungen verwendet. (Die Details würden hier zu weit führen; sie werden in [139, 140] erklärt.) Das führt in einzelnen Fällen dazu, daß Erweiterungsmengen im Sinne von Horty sich von solchen im Sinne der PDL (und *a fortiori* auch von solchen der DPDL) unterscheiden. Ein solcher Fall (*The Order Puzzle*) wird eingehend in [139, Kap. 4.2] und [140, Kap. 8.2] besprochen.

VIII.

ÜBERZEUGUNGSWANDEL
BELIEF REVISION

1. Epistemische Rechtfertigung

Zwei Rechtfertigungsprobleme. Karl sieht, wie sein Freund Peter in einem roten Auto vorüberfährt. Er glaubt, daß wenn Peter ein Auto steuert, dann gehört ihm das Auto. Karl glaubt daher:

(R) Peter besitzt ein rotes Auto.

Die Frage stellt sich, ob diese Überzeugung Karls gerechtfertigt ist. Nun folgt R im Beispiel aus weiteren Aussagen, von denen wir annehmen, daß sie zu Karls Überzeugungen gehören, nämlich diesen:

(P) Peter sitzt am Steuer eines roten Autos.
(Q) Wenn Peter am Steuer eines Autos sitzt, dann gehört das Auto ihm.

Die Überzeugung R ist eine deduktive Folge aus P und Q. Wenn wir davon ausgehen, daß deduktive Folgerung die Eigenschaft gerechtfertigt zu sein von den Prämissen an die Konklusion überträgt, dann werden wir so verwiesen auf die Frage, ob diese Prämissen gerechtfertigt sind. Darf Karl aufgrund seiner Wahrnehmung davon ausgehen, daß Peter im Auto sitzt und dieses wirklich rot ist? Wie zuverlässig ist das Konditional Q? Vielleicht ist es besser, die Rechtfertigungsfrage nicht (fundamentierend) für jede Überzeugung einzeln, sondern (kohärentistisch) für die Gesamtheit von Karls Überzeugungen zu stellen. Passen sie gut zusammen? Geben sie eine gute Basis für erfolgreiches Handeln ab? Wenn das so ist, dann könnten wir in einem sekundären Sinne sagen, daß Karls Überzeugungen gerechtfertigt sind. Wie dem auch sei, wir befinden uns hier in der Ausgangsstellung der traditionellen Erkenntnistheorie: der Betrachtung eines Erkenntnissubjektes zu einem bestimmten Zeitpunkt und der Frage, ob das Subjekt zu diesem Zeitpunkt zurecht glaubt, was es glaubt. Das ist das Problem *statischer Rechtfertigung.*

Neben dieser traditionellen Fragestellung gibt es eine andere, ebenso naheliegende, in der es um die Rechtfertigung von Überzeugungen geht – oder besser: des Übergangs von einem Überzeugungssystem zu einem anderen. Im vorgestellten Szenario nehmen wir an, daß Karl erst P und Q, und also R glaubt, aber nun aus zuverlässiger Quelle erfährt, daß Peters Auto grün, also sicher nicht rot ist. Er erfährt also etwas, das $\neg R$ impliziert und so seinen bisherigen Überzeugungen – bezeichnen wir sie mit K – widerspricht. Karl muß seine Überzeugungen ändern. Es ist klar, daß er die Überzeugung, daß Peters Auto grün ist, nur annehmen kann, wenn die widersprechende Überzeugung R aus dem Weg geräumt ist. Karl muß also zu einem Überzeugungssystem übergehen, welches wir mit $K - R$ ("K ohne R") notieren wollen. Aber aufgrund der deduktiven Vernetzung mit P und Q gibt es hier mindestens drei Möglichkeiten:

(a) $K - R$ hält an P fest und gibt Q auf:
(b) $K - R$ hält an Q fest und gibt P auf;
(c) $K - R$ gibt P und Q auf.

Jede dieser Möglichkeiten kann sich in weitere verzweigen, je nachdem, ob P oder Q selbst wiederum aus weiteren Überzeugungen in K folgen. Im Extremfall könnte Karl in kartesische Meditationen verfallen und so ziemlich alle seiner Überzeugungen aufgeben. Karl wird sich entscheiden müssen und die Frage lautet nun: Was rechtfertigt seine Entscheidung? Wir können dies das Problem der *dynamischen Rechtfertigung* nennen.

Reduktion? Es ist merkwürdig, wie sehr die traditionelle Erkenntnistheorie auf das Problem der statischen Rechtfertigung fixiert ist und das andere Problem beinahe völlig ausblendet, obgleich dieses doch im Alltag das bei weitem drängendere ist. Denn nur gelegentlich sind wir herausgefordert, vor uns zu rechtfertigen, wovon wir bereits überzeugt sind; beinahe ständig jedoch, stehen wir vor dem Problem, auf Neues mit der Weiterentwicklung unserer Überzeugungen in die eine oder andere Richtung zu reagieren. Nur in der Tradition des Pragmatismus – in der *theory of inquiry* bei Peirce, James und Dewey – findet sich eine systematische Auseinandersetzung mit dem Problem dynamischer Rechtfertigung.

Der blinde Fleck könnte Ergebnis eines verführerischen Bildes sein. Danach kommen wir zu unseren Überzeugungen aufgrund kausaler Prozesse.

Das Überzeugungssystem eines Subjektes (zu einem bestimmten Zeitpunkt) ist die aggregierte Wirkung solcher Prozesse, die ihre Ursachen außerhalb des Subjektes haben. Da das Subjekt diese Prozesse nicht steuern kann, stellt sich die Frage gar nicht, inwiefern das Subjekt dabei Normen der Vernunft folgt. Allenfalls das Resultat läßt sich an der Absicht messen, ein wahres Bild der Welt abzugeben.[1] Aber schon eine oberflächliche Betrachtung von einfachen Episoden der Überzeugungsänderung (wie das oben beschriebene Beispiel) zeigt, daß das Bild nicht richtig sein kann oder zumindest gravierend unvollständig ist. Offensichtlich sind die Optionen in solchen Episoden selten derart, daß man eine Wahl auch einem blinden kausalen Prozess – etwa Würfeln – überlassen wollte. Dem Erwürfeln einer Option ziehen wir beinahe immer vernünftiges Nachdenken vor. Wir wägen die Optionen gegeneinander ab in der Erwartung, eine gute Wahl treffen zu können. Was "vernünftiges Nachdenken" und "gute Wahl" in diesem Zusammenhang bedeuten, ist der Gegenstand dieses Kapitels.

Gibt es einen besseren Grund, dynamische Rechtfertigung auszublenden und statische Rechtfertigung zum Angelpunkt der Erkenntnistheorie zu machen? Ein solcher Grund könnte die Möglichkeit sein, das Problem dynamischer Rechtfertigung auf das statischer Rechtfertigung zurückzuführen. Wir wollen jetzt kurz betrachten, wie solche Möglichkeiten im Prinzip aussehen könnten und wie es um ihre Aussichten auf Erfolg steht.

Was immer statische Rechtfertigung ($Sger$) sein mag, ihr Objekt ist ein Überzeugungszustand K:

$$Sger(K).$$

Dagegen geht es im dynamischen Fall ($Dger$) um die Rechtfertigung eines Übergangs von einem Überzeugungszustand K zu einem weiteren K'. D.h. das Objekt der Rechtfertigung ist ein Paar von Überzeugungszuständen, schematisch:

$$Dger(K, K').$$

Wie könnten wir den dynamischen Fall auf den statischen Fall reduzieren?

In einem ersten Versuch könnten wir einfach sagen: Der Übergang von K zu K' ist D-gerechtfertigt, wenn K' S-gerechtfertigt ist – also:

1. $Dger(K, K')$ gdw $Sger(K')$.

Der Ausgangszustand K taucht rechts gar nicht auf. Danach wäre der Übergang von einem beliebigen Zustand K zu einem S-gerechtfertigten

[1] Kuhns Untersuchungen und Thesen zu rational nicht rekonstruierbaren Paradigmenwechseln in der Wissenschaft haben sicher auch dazu beigetragen, das Problem der dynamischen Rechtfertigung einer normativen Behandlung zu entziehen und es in die Wissenschaftsgeschichte auszulagern.

Zustand K' immer vernünftig (D-gerechtfertigt). Das klingt merkwürdig, weil es dem Begriff der statischen Rechtfertigung den Boden zu entziehen droht. Denn auch statische Rechtfertigung ist ein relationaler Begriff: Es muß Ressourcen der Rechtfertigung geben. Der jetzt besprochene Versuch schließt aber aus, daß der Ausgangszustand K des Subjekts als Resource in irgendeinem Sinne infrage kommt: weder (internalistisch) als dem Subjekt verfügbare Gründe, noch (externalistisch) als Träger irgendwelcher epistemisch relevanter Eigenschaften. So bleibt völlig unklar, welches die Ressourcen statischer Rechtfertigung sein sollen.

Wir könnten versuchen, den Ausgangszustand K auf der rechten Seite so im Spiel zu halten:

2. $Dger(K, K')$ gdw $Sger(K)$ zum Zeitpunkt t und $Sger(K')$ zum Zeitpunkt t' des Übergangs.

Aber daraus folgt, daß es aus einem zum Zeitpunkt t epistemisch schlechten Anfangszustand K keinen gerechtfertigten Übergang zu irgendwelchen Überzeugungen K' gibt. S-ungerechtfertigte Zustände wären in diesem Sinne ausweglos. Und es scheint ebensowenig plausibel, daß es zu schlechten Überzeugungen K' keinen guten Weg gibt. Wir können ja auch auf falsche Konklusionen auf gültige Weise schließen.

Vielleicht geht die Sache besser, wenn wir S-Rechtfertigung als einen gradierbaren Begriff auffassen und eine entsprechende Relation einführen. Dann könnten wir sagen:

3. $Dger(K, K')$ gdw K' S-gerechtfertigter ("besser") als K ist.

(Man könnte auch etwas schwächer fordern, daß K nicht besser als K' sei.)

Jetzt ist ausgeschlossen, daß es schlechte Übergänge zu besseren Überzeugungen gibt. Aber könnte K' nicht S-besser als K sein, ohne daß sich der Übergang rechtfertigen läßt? Hier ist ein einfaches Beispielmuster: Sei A eine Überzeugung in K; K' enthalte genau diejenigen Überzeugungen in K, die mindestens so S-gut wie A sind. K' repräsentiert also einen "skeptischen Rückzug" von K mit A als Maßstab. Wenn K' echt in K enthalten ist, dann ist K' sicher S-besser als K. Aber da wir A beliebig wählen können, ist der durch den Rückzug bedingte Informationsverlust willkürlich und daher nicht D-gerechtfertigt.

Ebenso schließt 3 gute Übergänge zu S-schlechteren Überzeugungen aus. Wenn wir die gerade angestellte Überlegung "dualisieren", sehen wir, daß das genausowenig plausibel ist. Jede (nicht-deduktive) Erweiterung K' von K erhöht das Risiko des Irrtums. Wenn K' ein höheres Irrtumsrisiko als K trägt, dann kann K' nicht besser S-gerechtfertigt sein als K. Also sind riskante Erweiterungen eines Überzeugungszustandes niemals D-gerechtfertigt. Aber kann es nicht mitunter vernünftig sein, Überzeugungen riskant zu erweitern? Falls nicht, müßten wir uns auf ein absolut sicheres Korpus ("*cogito*") zurückziehen.

Es scheint wenig aussichtsreich zu sein, dynamische auf statische Rechtfertigung zurückführen zu wollen. Vielleicht aber läßt sich umgekehrt statische Rechtfertigung als Derivat dynamischer Rechtfertigung auffassen? Etwa so:

4. Ein Überzeugungszustand ist S-gerechtfertigt, wenn er als Endpunkt einer Reihe D-gerechtfertigter Überzeugungsänderungen darstellbar ist.

Es stellt sich sofort die Frage nach der Rechtfertigung des Anfangszustandes in der hypothetischen Reihe. Ein Regreß nach bekanntem Muster droht. Wenn wir einen Anfangszustand in relevanter Weise auszeichnen wollen, dann kann dynamische Rechtfertigung dabei jedenfalls keine Rolle spielen. Die Auszeichnung wird plausiblerweise eine Form von S-Rechtfertigung sein, die nicht auf D-Rechtfertigung zurückgeführt werden kann. Alternativ könnten wir auf die Auszeichnung von Anfangszuständen verzichten und diese als wesentliche Parameter von D-Rechtfertigung auffassen. Nun stellt sich die Frage, ob ein so relativierter Begriff von S-Rechtfertigung die theoretischen Rollen übernehmen kann, die ihm in der Erkenntnistheorie zugedacht sind.

An dieser Stelle brechen wir die Betrachtung von Reduktionsstrategien ab – denn als Ergebnis dürfen wir festhalten, woran uns hier liegt: Wie wir die Sache auch drehen und wenden, dynamische Rechtfertigung stellt ein eigenständiges erkenntnistheoretisches Problem dar.

2. Überzeugungszustände

Wir wollen das Problem in seiner einfachsten, allgemeinsten Form angehen: Information i wirkt auf einen Überzeugungszustand K ein, worauf ein Nachfolgezustand K' erzeugt wird.

$$\begin{tikzcd}
& i\colon A \arrow[dl, rightsquigarrow] & \\
K \arrow[rr] & & K'
\end{tikzcd}$$

K' muß nicht von K verschieden sein, denn i könnte in K schon berücksichtigt sein, oder K könnte so sein, daß i keinen Überzeugungswandel auslöst. Über die Natur von i werden wir hier nichts weiter sagen müssen. Wir betrachten i einfach funktional als Stimulus für einen Überzeugungswandel. Wichtig ist nur, daß ein solcher Stimulus einen kognitiven Gehalt hat, der in einer Aussage A gefaßt werden kann. Eigene Wahrnehmungen, eigene Überlegungen oder Aussagen anderer kommen typischerweise als Stimuli in Betracht.

Überzeugungszustände müssen wir etwas näher betrachten. Die Art und Weise, wie wir diese darstellen und welche Bedingungen sie dabei erfüllen sollen, wird wesentlich für die angestrebte Theorie des Überzeugungswandels sein. Überzeugungszustände werden wir hier durch Mengen von Aussagen in einer vorgegebenen Sprache darstellen. (Wir werden später sehen, daß wir noch eine weitere Strukturkomponente benötigen.) Die Theorie soll normativ sein in dem Sinne, daß ein Nachfolgezustand K' die beste, die vernünftigerweise gebotene Wahl aus einer Reihe von Optionen ist. Was immer "beste" etc. hier genau heißen mag, so können wir doch schon fragen, ob ein inkonsistenter Überzeugungszustand ein Optimum in diesem Sinne sein kann? Sicher nicht, wenn es konsistente Alternativen gibt. Könnte ein Zustand, der nicht deduktiv abgeschlossen ist, optimal sein? Wohl kaum, denn ein solcher Zustand würde auf Information verzichten, die risikolos zu haben ist. Ein Nachfolgezustand K' sollte daher konsistent und deduktiv abgeschlossen sein. Ein solcher Zustand K' ist aber Ausgangspunkt für weitere Überzeugungsänderungen. Daher legen wir unsere Theorie am besten so an, daß sie sich auf Änderungen im Bereich *guter* Überzeugungszustände beschränkt, d.h. solcher Zustände K, welche diese zwei Bedingungen erfüllen:

- K ist *konsistent*, d.h. $\perp \notin K$, und
- K ist *deduktiv abgeschlossen*, d.h. $K = \mathrm{Cn}(K)$.

Tatsächlich wird unsere Theorie etwas allgemeiner ausfallen: Wir werden weder ausschließen, daß ein Überzeugungszustand um die Kontradiktionsaussage $\perp$ zu einem inkonsistenten Zustand erweitert wird, noch daß Änderungsoperationen für den Grenzfall FML, der Menge aller Sätze einer Sprache, definiert sind. Natürlich könnten wir entsprechende Beschränkungen einführen. Der zusätzliche Aufwand würde jedoch ohne echten Ertrag bleiben.

Die Überzeugungsmengen realer Person sind sicher nicht deduktiv abgeschlossen. Gefährdet das die Anwendbarkeit der Theorie? Nicht, wenn wir die intendierte Anwendung richtig verstehen. Die manifesten Überzeugungen echter Personen sind endlich und also nicht deduktiv abgeschlossen. Aber manifeste Überzeugungen *verpflichten* auch echte Personen zur Anerkennung weiterer Überzeugungen. Deduktiv abgeschlossene Überzeugungsmengen sind also solche, zu denen echte Personen verpflichtet sind, auch wenn die Anerkennungsperformanz solcher Personen hinter ihren Verpflichtungen zurückbleibt, ja zurückbleiben muß. In der Anwendung fragen wir daher nach legitimen Übergängen von einer Menge von Überzeugungen, zu denen echte Personen verpflichtet sind, zu einer anderen.

Die Gesamheit der Überzeugungen einer realen Person sind im allgemeinen auch nicht konsistent. Reale Personen überschauen nicht die Gesamtheit der logischen Konsequenzen ihrer expliziten Überzeugungen. So kann ihnen insbesondere auch entgehen, daß sie sich zu einer widersprüchlichen Überzeugungsmenge verpflichtet haben. Das kann daran liegen, daß reale Personen in sehr unterschiedlichen Kontexten agieren und Überzeugungen ausbilden. So kann eine Person in einem Kontext zur Überzeugung A, in einem anderen Kontext zu $\neg A$ gelangen – im expliziten oder im Verpflichtungssinne. Dabei mag es jedoch keinen Kontext geben, in dem die Person jemals beide Überzeugungen aktivieren müßte. In solchen Fällen können wir sagen, daß die Person, je nach Kontext, mit zwei konsistenten Überzeugungssystemen arbeitet. Es gibt keinen zwingenden Grund, Überzeugungssysteme entlang Körpergrenzen zu individuieren. Auch Gruppen echter Personen können ein gemeinsames Überzeugungssystem hinsichtlich bestimmter Fragen haben. Für unsere Zwecke ist es besser, Überzeugungssysteme nicht wie Dinge, sondern wie Eigenschaften aufzufassen: Eine Person kann mehrere solcher Eigenschaften auf sich vereinen, und verschiedene Personen können solche Eigenschaften teilen.

Was, wenn eine reale Person widersprüchliche Überzeugungen zu *einer* Fragestellung, in *einem*, nicht sinnvoll teilbaren Kontext hat? Auch das kommt vor, da Überzeugungen versteckte Widersprüche enthalten können; man denke etwa an Naive Mengenlehre oder an Freges Theorie der Arithmetik in seinen *Grundgesetzen*. Über die Änderung solcher Überzeugungsmengen hat unsere Theorie wenig Interessantes zu sagen. Der Ausweg ist *common sense*: Sobald wir der Situation gewahr werden, nehmen wir Teile unserer endlichen Repräsentation der inkonsistenten Menge zurück und zeigen oder hoffen, daß unsere Überzeugungen nun widerspruchsfrei sind. Ein solches Zurückbauen einer endlichen Menge von Überzeugungen, die einen Widerspruch implizieren, ist alles andere als trivial und kann ebenfalls zum Gegenstand einer Theorie gemacht werden. Einen Beitrag dazu liefert die Theorie der Überzeugungsänderung *via* endlicher Basen, auf die wir später kurz zurückkommen werden (p. 500).

So wie wir den Begriff hier verstehen wollen, sind Überzeugungssysteme – in logischer Sprechweise – *Theorien* in einer aussagenlogischen Sprache. Theorien sind Mengen von Sätzen, die unter einer Operation Cn abgeschlossen sind, die wir gewöhnlich als logische Konsequenz interpretieren. Unter logischer Konsequenz dürfen wir uns hier die bekannte Relation der Ableitbarkeit aus Annahmen in der klassischen Logik vorstellen. Bis auf wenige Stellen kann die Theorie jedoch ohne spezifische Annahmen über die zugrundeliegende Logik entwickelt werden. Es reichen ganz allgemeine Eigenschaften der Ableitbarkeitsrelation, die sich in Eigenschaften der Operation Cn übersetzen. Es sei hier an die charakterischen Eigenschaften einer Konsequenzoperation erinnert.

Inklusion:	$X \subseteq \mathrm{Cn}(X)$;
Monotonie:	wenn $X \subseteq Y$, dann $\mathrm{Cn}(X) \subseteq \mathrm{Cn}(Y)$;
Idempotenz:	$\mathrm{Cn}(\mathrm{Cn}(X)) = \mathrm{Cn}(X)$;
Kompaktheit:	wenn $A \in \mathrm{Cn}(X)$, dann $\exists X' \subseteq X :$ X' ist endlich und $A \in \mathrm{Cn}(X')$.

Wir schreiben manchmal auch relational

$$X \vdash A \text{ für } A \in \mathrm{Cn}(X) \quad \text{und} \quad A \equiv B \text{ für } \mathrm{Cn}(A) = \mathrm{Cn}(B).$$

Daß ein Überzeugungszustand K unter logischer Konsequenz abgeschlossen sein soll, bedeutet also in relationaler Notation, daß

$$A \in K \text{ gdw } K \vdash A.$$

3. Idee und Anwendungen

In diesem Abschnitt werden wir anhand dreier Anwendungen die wesentlichen Ideen der sogenannten AGM-Theorie der Überzeugungsänderung vorstellen. "AGM" steht kurz für drei Autoren, welche die Theorie, in wechselnder Kooperation, maßgeblich entwickelt haben: der argentinische Rechtstheoretiker Carlos Alchourrón, der schwedische Philososoph Peter Gärdenfors und der australische Logiker David Makinson. Die AGM-Theorie besteht aus den AGM-Postulaten für die Operationen der Kontraktion und Revision einerseits, sowie einer Reihe dazu passender Konstruktionen andererseits. Die hier besprochenen Anwendungen sind zugleich die wichtigsten Inspirationsquellen für die Theorie. Die zuerst behandelte "Anwendung" ist tatsächlich eine umfassende und detaillierte Erkenntnistheorie, die das Lebenswerk des amerikanischen Philosophen Isaac Levi darstellt und in seinen wesentlichen Zügen den Arbeiten von AGM vorausging.

Gegeben ein Überzeugungszustand K, so gibt es genau drei Möglichkeiten wie sich ein Folgezustand K' zu K im Mengenvergleich verhalten kann:

- K ist in K' enthalten. Dann ist K' eine Erweiterung von K.
- K' ist in K enthalten. Dann ist K' eine Verkleinerung von K.

– Weder ist K in K' noch umgekehrt K' in K enthalten. Dann ist K' eine Abänderung von K.

Betrachten wir eine Aussage A als Auslöser einer Änderung von K, dann gibt es entsprechend drei Arten von Operationen f, die Überzeugungszustände anhand eines Aussagenparameters in neue Überzeugungszustände zu überführen.

– Im Erweiterungsfall $K \subseteq f(K, A)$ ist $f(K, A)$ eine *Expansion* von K um A; fortan $K + A$ geschrieben.
– Im Verkleinerungsfall $f(K, A) \subseteq K$ ist $f(K, A)$ eine *Kontraktion* von K um A; fortan $K - A$.
– Im Abänderungsfall ist *f(K,A)* eine *Revision* von K um A; fortan $K * A$.

Um die Gedanken in die richtige – und natürliche – Richtung zu lenken, wollen wir jetzt schon annehmen, daß das Resultat einer Expansion oder Revision um eine Aussage diese Aussage enthalten sollen, während Kontraktionen den Aussageparameter aus dem Anfangszustand entfernen. Dabei sind harmlose Grenzfälle zugelassen, in denen diese Operationen leerlaufen: Kontraktionen, die im Anfangszustand gar nicht vorhandene Aussagen “entfernen”, und Expansionen bzw. Revisionen, die “hinzufügen”, was schon vorhanden ist. Bevor wir diese und andere Bedingungen genauer besprechen, wollen wir beispielhaft umreißen, zu welchen Fragen und Problemen eine Theorie solcher Operationen einen Beitrag leisten könnte.

Levi: Erkenntnistheorie als kognitive Entscheidungstheorie. Unsere Überzeugungen sind Überzeugungen darüber, was wahr ist. Wir entwickeln unsere Überzeugungen, in der Absicht, der Wahrheit möglichst nahe zu kommen. Je mehr wir irren, umso weiter entfernen wir uns von diesem Ziel; je mehr Wahrheiten unter unseren Überzeugungen sind, umso näher kommen wir ihm. Diese beiden Aspekte ziehen leider in entgegengesetzte Richtungen. Einerseits minimieren wir Irrtum, indem wir möglichst wenig glauben; andererseits maximieren wir wahre Überzeugungen, indem wir möglichst viel glauben. Am wenigsten riskieren wir zu irren im Zustand $\mathrm{Cn}(\emptyset)$, welcher nur die logischen Wahrheiten enthält; am wenigsten riskieren wir, eine Wahrheit zu verpassen im Zustand $\mathrm{Cn}(\perp)$ [= FML], in dem wir alles glauben. Das Ziel kann also nur darin bestehen, ein vernünftiges Gleichgewicht zwischen der Vermeidung von Irrtum und der Suche nach Wahrheit herzustellen. Wissen ist nicht risikofrei zu haben und wir müssen uns entscheiden, welcher Informationsgewinn uns welches Risiko wert sein sollte.

Dieses Entscheidungsproblem wird schon in unserem kleinen Anfangsbeispiel deutlich. Peter sieht sich genötigt, die Überzeugung, daß Karl ein rotes Auto besitzt (R) aufzugeben. Auf p. 478 haben wir mindestens drei Optionen (a-c) identifiziert, wie Peter seinen Irrtum korrigieren kann. In jedem dieser drei Kandidaten für einen Nachfolgezustand $K - R$ verliert Peter weitere und jeweils verschiedene Information. Peter muß sich entscheiden. Soll diese Entscheidung nicht willkürlich sein, so muß der Überzeugungszustand K, in dem Peter sich befindet, über Ressourcen verfügen, die eine *Auswahl* in (a–c) begründen können. Auch wenn wir hier meist K als eine Variable für Mengen von Aussagen behandeln, so ist doch klar, daß Überzeugungssysteme nicht einfach als solche Mengen aufgefaßt werden können, wenn wir das theoretische Interesse auf die vernunftgeleitete Dynamik von Überzeugungssystemen richten. Überzeugungssysteme haben eine nicht bloß deduktive, sondern auch präferenzielle Struktur – ein Aspekt, der, wie wir später sehen werden, wesentlich für die Modellierung von Überzeugungsänderung sein wird.

Das ist eine zwingende und daher unkontroverse Einsicht, die am Anfang jeder Theorie vernünftiger Überzeugungsänderung steht. An diese Einsicht schließen sich unmittelbar Fragen an, mit deren Beantwortung man substanzielle Positionen in der Erkenntnistheorie bezieht. Wie ist der nötige Auswahlmechanismus, mit dem das Entscheidungsproblem gelöst wird, zu interpretieren? Grob gesprochen, kommen hier nur zwei Interpretationen infrage: Wahrscheinlichkeit oder Nutzen (beides in einem weiten Sinne). Nach der ersten Interpretation wäre – etwas vereinfachend – eine der drei Optionen die wahrscheinlichste. Dann müßten wir annehmen, daß eine Theorie vernünftiger Überzeugungsänderung erfordert, daß Peters Überzeugungen wahrscheinlichkeitsgradiert oder jedenfalls wahrscheinlichkeitsgeordnet sind. Isaac Levi wendet sich mit starken Argumenten gegen diese Auffassung und besteht auf der Nutzeninterpretation. Die Optionen haben einen je verschiedenen Informationswert ("Nutzen") und die Auswahl geschieht mit dem Ziel, in einen Nachfolgezustand mit möglichst wertvoller Information zu wechseln. Es mag sein, daß Informations- und Wahrscheinlichkeitswerte (oder -ordnungen) die gleiche formale Struktur haben (in Levis Theorie ist das so). Aber das sollte nicht dazu verleiten, beides in eins zu setzen.

Eine zweite Frage betrifft die Rechtfertigung von Überzeugungsänderungen überhaupt. Bisher sind wir von einer Situation ausgegangen, in der ein Entschluß zu einer Überzeugungsänderung – Expansion, Kontraktion oder Revision – gefaßt wurde und wir haben danach gefragt, wie eine solche Änderung auf legitime Weise ins Werk gesetzt werden kann. Aber wie kann der Entschluß selbst gerechtfertigt werden? Unter welchen Umständen sind Expansionen, Kontraktionen oder Revisionen legitim? Die Frage mag zunächst befremden, da wir doch ständig unsere Überzeugungen in der einen oder anderen Weise ändern und diese Praxis doch sicher nicht *tout court* problematisch sein kann. Aber die Problematik erschließt sich schnell, wenn wir uns an das Doppelziel "Meide Irrtum, suche Wahrheit!" (W. James) erinnern. Wenn wir expandieren, dann gehen wir ein Irrtumsrisiko ein, zu dem wir im Ausgangszustand K nicht bereit sind; wenn wir kontrahieren, verlieren wir Information, die wir in K für wichtig und wahr halten; wenn wir revidieren, dann tun wir beides. [2] Auch diese zweite Frage ist Gegenstand von Levis Erkenntnistheorie.

Legitime Expansionen sind von zweierlei Art. Zunächst gibt es Routinen, mit denen wir unser Überzeugungssystem normalerweise erfolgreich mit der Welt abgleichen. Dazu gehört die Wahrnehmung und die Übernahme dessen, was uns aus vertauenswürdiger Quelle mitgeteilt wird (beides präsent in unserem kleinen Ausgangsbeispiel). Im Falle der Wahrnehmung ist es uns schlechthin unmöglich, die Routine zu verweigern. Im Falle des Zeugnisses anderer haben wir es mit einer vorgängigen Entscheidung zu tun, der Routine zu folgen. Im Gegensatz zur Wahrnehmung ist diese Entscheidung rücknehmbar. Aber so lange die Entscheidung in Kraft ist, führt sie zu routinemäßigen Expansionen. Ist das Vertrauen auf Wahrnehmung und qualifizierte Zeugen unter normalen Umständen legitim? Unter der Voraussetzung, daß wir skeptischen Argumenten nicht folgen müssen, ist das sicher so.

Die andere Art legitimer Expansion verdankt sich der Notwendigkeit, Fragen auch dann beantworten zu müssen, wenn der gegenwärtige Stand der Überzeugungen nicht deduktiv hinreichend für eine Antwort ist. Am besten wäre eine möglichst informative Antwort. Aber da wir uns gegen Irrtum schützen sollten, spielt nun auch die Frage eine Rolle, wie wahrscheinlich die möglichen Antworten sind. Wir benötigen also eine Verteilung von epistemischem Nutzen und Wahrscheinlichkeit über das Menü möglicher Antworten. Eine Abwägung dieser Werte, zusammen mit einer Schwelle persönlicher Risikobereitschaft, wird eine – möglicherweise disjunktive – Antwort als die beste identifizieren um die wir nun expandieren sollten. Solche Überzeugungsänderungen nennt Levi induktive Expansionen. Sie sind legitim relativ zu einer Partition möglicher Antworten auf eine Frage, einer Verteilung von Nutz- und Wahrscheinlichkeitswerten über diese Partition und einer gegebenen Bereitschaft zum Risiko.

Nun zur Frage der legitimen Kontraktionen. Es ist klar, daß Routineexpansionen prinzipiell "ohne Gewähr" sind. Insbesondere bieten sie keine Gewähr, daß das die unwillkürlich aufgenommene Information logisch verträglich mit den Ausgangsüberzeugungen ist. So kann durch Routineex-

[2] Siehe dazu auch den Aufsatz [167] von Kripke.

pansion ein inkonsistenter Zustand entstehen. Das erzwingt, so Levi, eine anschließende Kontraktion. Dagegen ist sicher nichts zu sagen: So gehen wir tatsächlich und richtig vor. Es ist nur so, daß die Theorie mit solchen Rückzügen aus widersprüchlichen Überzeugungen prinzipiell überfordert ist. Denn Überzeugungszustände sollen ja deduktiv abgeschlossen sein. Alle inkonsistenten Überzeugungszustände legen daher dieselbe triviale Überzeugungsmenge fest. Inkonsistente Zustände können also nicht nach ihren jeweiligen Überzeugungsmengen differenziert werden. So bleibt offen, wie Überzeugungssysteme, deren Überzeugungsmenge die Menge aller Aussagen ist, so individuiert werden können, daß die Art von Kontraktion die hier im Blick ist, gelingen kann.[3]

Es gibt einen weiteren Grund, Überzeugungen manchmal zurückzunehmen. Wenn wir glauben, daß A der Fall ist, dann ist die konkurrierende Hypothese $\neg A$ ausgeschlossen. Wir können sie gar nicht sinnvoll betrachten, d.h. z.B. prüfen, welchen Beitrag sie, vor dem Hintergrund unserer anderen Überzeugungen, zur Beantwortung von Fragen leisten könnte. Eine vorurteilsfreie Betrachtung von $\neg A$ setzt voraus, daß wir versuchsweise A außer Kraft setzen, d.h. in den Zustand $K - A$ wechseln. Wenn wir konkurrierende Hypothesen als Antworten auf Fragen ernsthaft erwägen wollen, dann müssen solche öffnenden Kontraktionen gerechtfertigt sein.

Wenn dies die einzigen Arten sind, wie wir den Zuwachs von Risiko (in Expansionen) und den Verzicht auf Information (in Kontraktionen) rechtfertigen können, dann ergeben sich daraus auch die Bedingungen, unter denen wir zurecht eine Revision unserer Überzeugungen vornehmen können. Denn diese erhöhen ja typischerweise zugleich das Irrtumsrisiko (wie in Expansionen) und geben Information auf (wie in Kontraktionen). Legitime Revisionen können also nur eine legitime Expansion mit einer legitimen Kontraktion kombinieren. Die richtige Abfolge ist diese: Auf eine Kontraktion $K - \neg A$ mit der Absicht, K für A als Antwortmöglichkeit auf eine Frage zu öffnen, folgt die (induktive) Expansion um A. Die umgekehrte Reihenfolge, $K + A - \neg A$, lädt – im Hauptfall $\neg A \in K$ – offensichtlich nicht nur das bereits besprochene Problem ein, wie man sich durch Kontraktion aus dem Sumpf des Widerspruchs befreien kann. Sie ist schon aus dem Grunde nicht gangbar, weil in diesem Fall $K + A$ keine gerechtfertigte induktive Expansion sein kann; denn aus der Perspektive von K stellt A keine mögliche Antwort auf irgendeine Frage dar.

Wir haben hier Levis Theorie etwas ausführlicher besprochen – wenn auch nur in den Grundzügen skizziert –, weil sie zeigt, wie natürlich sich an

[3] Es liegt nahe, auf die Bedingung des deduktiven Abschlusses zu verzichten, wie das in der bereits erwähnten Theorie der Basisänderungen geschieht; cf. p. 500.

die im Kern unkontroverse, rein formale Theorie des Überzeugungswandels streitbare philosophische Theorenbildung anschließen kann, und wie die formale Theorie selbst weitere Probleme aufdeckt (und den Bereich möglicher Lösungen absteckt), die wesentlich zur modernen Erkenntnistheorie gehören.

Gärdenfors: Epistemische Semantik von Konditionalsätzen. In Levis Theorie stehen Expansionen und Kontraktionen im Vordergrund. Jede Revision sollte als Verkettung einer Kontraktion mit einer Expansion darstellbar sein. Expansionen sind – wie wir gleich etwas genauer erklären werden – einfache logische Operationen: Mengenerweiterungen, die unter logischer Konsequenz abgeschlossen werden. So besteht der Kern einer Theorie von Überzeugungsänderungen in einer Theorie der Kontraktionen.

Auch in der AGM-Theorie ist das so. Aber in einer wichtigen Quelle und Anwendung der Theorie, nämlich Peter Gärdenfors' [96] epistemischer Semantik bestimmter Konditionalsätze, werden allein Revisionen von Überzeugungszuständen betrachtet. Als Ausgangspunkt dieser Anwendung wird oft Ramsey [239] zitiert:

> If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q.

Hier wird eine Möglichkeit angedeutet, wie man sich Klarheit über die Bedeutung von Konditionalen verschaffen kann: Ramsey schlägt eine Akzeptanzbedingung für solche Aussagen vor. Nun sind Akzeptanzbedingungen keine Wahrheitsbedingungen und man könnte der Auffassung sein, daß nur letztere in der Lage sind Bedeutungen anzugeben. Aber so allgemein formuliert, ist das sicher nicht richtig. Es gibt viele Arten von Aussagen, von denen wir nicht sagen würden, daß sie wahr oder falsch sind, sondern daß ein Sprecher sie unter bestimmten Umständen zurecht behauptet oder bestreitet. Typische Aussagen dieser Art sind Geschmacksurteile. In diesen Fällen geben die Akzeptanzbedingungen an, was ein Sprecher mit einer solchen Aussage ausdrücken möchte. Auch bestimmte Konditionale könnten von dieser Art sein. Im Kapitel IV über Konditionallogik haben wir diese Position ausführlich behandelt.

Gärdenfors' Theorie basiert auf einem einfachen Vorschlag, wie wir Äußerungen von Konditionalen wie "Wenn Karl nicht gebremst hätte, wäre er verunglückt" verstehen sollten. Der Sprecher drückt aus: Wenn ich zu meinen gegenwärtigen Überzeugungen hinzunehme, daß Karl nicht gebremst hat, dann bin ich in einem Zustand, in dem ich glaube, daß er verunglückt ist. Das Konditional ist also eine Aussage nicht – wie bei Lewis

– über die Eigenschaften bestimmter möglicher Welten, sondern über die Eigenschaften eines potentiellen Überzeugungszustandes. Damit hat die Theorie zumindest den Vorteil, daß sie keine Fragen über die Existenz alternativer Welten und unseren epistemischen Zugang zu ihnen aufwirft.

Ein Vergleich zwischen Gärdenfors' und Lewis' Ansätzen ist instruktiv. Um die Sache besser auf den Punkt zu bringen, wählen wir hier die stärkere Variante von Stalnaker.[4] (Der Junktor $\sqsupset$ möge für das kontrafaktische Konditional stehen.)

Lewis (eigentlich Stalnaker): $A \sqsupset B$ ist genau dann wahr in Welt a, wenn B wahr ist in derjenigen Welt b, in der auf a-naheliegendste Weise A der Fall ist.

Gärdenfors: $A \sqsupset B$ wird genau dann in einem Überzeugungszustand K akzeptiert, wenn B akzeptiert wird im Zustand K', der aus K durch Hinzufügung von A entsteht.

Die Ähnlichkeit fällt sogleich auf. Wenn wir Wahrheit gegen Akzeptanz, Welt gegen Zustand, und Naheliegen gegen Hinzunehmen austauschen, dann können wir die Klauseln wechselseitig ineinander überführen. Klammern wir einmal Wahrheit/Akzeptanz und Welt/Zustand aus, so ist es offenbar die Rede vom "Naheliegen" bzw. vom "Hinzunehmen" (Ramsey: "adding"), welche hier die theoretische Last trägt und weiterer Klärung bedarf. Im Ansatz von Gärdenfors benötigen wir daher eine Theorie des Hinzunehmens! Wie wir sehen werden, stellte genau diese die AGM-Theorie mit ihrer Charakterisierung von Revisionen zu Verfügung. So bekommt die Klausel oben im Rahmen der AGM-Theorie einen genauen Sinn:

Gärdenfors (AGM): $A \sqsupset B \in K$ gdw $B \in K * A$.

Das ist der sogenannte *Ramsey-Test* für Konditionale. Leider stellt sich heraus, daß der Ramsey-Test nicht zur AGM-Theorie paßt: Nur in trivialen Fällen steht der Ramsey-Test nicht im Widerspruch zu den AGM-Postulaten. Mit Prinzipien wie dem Ramsey-Test – die in Überzeugungszustände Aussagen einführen, welche Änderungen solcher Zustände systematisch reflektieren – müssen wir also vorsichtig umgehen. Wir werden daher im Abschnitt über die Logik kontrafaktischer Konditonalsätze den (RT) in eine etwas schwächere Variante der AGM-Theorie einbetten. Trotz dieser Einschränkung dürfen wir jetzt schon systematische Ähnlichkeiten in der

[4] Zur Erinnerung: Im Gegensatz zu Stalnaker postuliert Lewis nicht, daß es immer genau eine der Ausgangswelt a ähnlichste Welt gibt, in der das Antezedens des Konditionals wahr ist. Wenn es mehrere ähnlichste Welten gibt, dann fordert die Wahrheit des Konditionals in a, daß das Konsequens in allen diesen Welten wahr ist.

Modellierung von kontrafaktischen Konditionalen einerseits und Revisionsoperationen andererseits erwarten. Diese Ähnlichkeiten werden wir im Abschnitt über Sphärensysteme näher untersuchen.

Alchourrón und Makinson: Derogation von Normen. Genauso wie Gerichte höherer Instanz Urteile von Gerichten niederer Instanz aufheben können, so gibt es im Gesetzeswerk eines Landes auch eine Hierarchie von Bestimmungen. Ganz oben in der Hierarchie steht die Verfassung; andere Gesetze sind dieser untergeordnet. Wenn ein Gesetz einer Verfassungsnorm widerspricht – was durch den Obersten Gerichtshof festgestellt wird –, dann kommt es zu dem, was Juristen *Derogation* nennen: Die Verfassung hebt das Gesetz teilweise oder ganz (Abrogation) auf.

Derogation kann in allen Hierarchiestufen des Rechts vorkommen. Damit eine Derogation effektiv ist, genügt es meist nicht, von einer bestimmten Vorschrift zu sagen, sie sei nun aufgehoben. Denn es mag andere Vorschriften geben, welche die beanstandete Norm ebenfalls implizieren. Deshalb bedient sich eine Derogation per Gesetz oft der Formel: "Alle diesem Gesetz widersprechenden Bestimmungen in anderen Gesetzen treten mit In-Kraft-Treten dieses Gesetzes außer Kraft." Hinter der scheinbar einfachen Formel verbirgt sich ein Problem, das wir bereits kennengelernt haben. Angenommen, Gesetz C soll aufgehoben werden. Nun mag es Gesetze A und B geben, welche zusammen, jedoch nicht einzeln, Gesetz C implizieren. Dann ist A-und-B eine der beabsichtigten Derogation widersprechende Bestimmung. Treten nun A *und* B außer Kraft oder nur eines von beiden? Das Gebot der Rechtssicherheit fordert, daß es für solche Fälle ein allgemein anwendbares Verfahren gibt, um zu bestimmen, was denn nun gilt und was nicht. Wie läßt sich der formale Kern dieses Verfahrens beschreiben?

Diese Frage haben sich Carlos Alchourrón und David Makinson in [6] gestellt:

> Suppose that A is a set of regulations, y is some proposition that is implied by A, and that for some reason a legislative body wants to eliminate y. In such a situation, the body may decide to reject y, with the intention of thereby rejecting implicitly whatever in A implies y, retaining the remainder. This we shall call derogation. (P. 127)

Die Derogation von Normen stellt vor ein Problem, welches völlig analog ist zur Rücknahme von Überzeugungen. Es *ist* das gleiche Problem, nur in anderen Termini. Das liegt daran, daß Normen weitere Normen einschließen genau so, wie Überzeugungen weitere Überzeugungen einschließen. Normen sind wie Überzeugungen durch Abhängigkeiten vernetzt. Im Falle von Überzeugungen sind die Abhängigkeiten deduktiver Natur. Im Falle von Normen mag das letztlich auch so sein. Aber darauf müssen wir uns hier nicht festlegen. Um die Operation der Derogation im Sinne einer Kontraktion verstehen zu können, genügt es, daß Normen eine propositionale Struktur haben und daß die Relation der Abhängigkeit zwischen Normen die Struktur logischer Folgerung hat. So können wir Normen durch Aussagen repräsentieren (ohne zu behaupten, daß sie solche *sind*!) und ein Korpus von Normen als Theorie auffassen (vgl. p. 484) – mehr ist nicht nötig, um die Derogation von Normen in den Anwendungsbereich der AGM-Theorie zu bringen.

4. Die AGM-Postulate

Die AGM-Theorie geht das Problem der Überzeugungsänderung von zwei Seiten an. Einerseits sollen die Operationen der Kontraktion und Revision durch eine Reihe von Bedingungen (Postulaten) implizit definiert werden. Andererseits wollen wir explizit angeben, wie ein Anfangszustand in einen entsprechend geänderten Zustand überführt werden kann. Schließlich sollen sich Postulate und Konstruktionen gegenseitig stützen: Wenn eine Kontraktion $K - A$ die Postulate erfüllt, dann kann sie aus K wie beschrieben, konstruiert werden; und umgekehrt haben wir $K - A$ auf die beschriebene Weise konstruiert, dann erfüllt $K-A$ die postulierten Bedingungen. Ebenso für Revisionen. Wir werden eine Reihe von Konstruktionen kennenlernen, die alle zu den Postulaten in der gerade beschriebenen Weise passen. Die Postulate und verschiedenen Konstruktionen repräsentieren also allesamt dieselben Änderungsoperationen. Diese Repräsentationsresultate verleihen der AGM-Theorie eine bemerkenswerte Stabilität. In diesem Abschnitt wenden wir uns den Postulaten zu.

Expansion. Zunächst wollen wir kurz die einfache Operation der *Expansion* behandeln. Sei K ein Überzeugungszustand und A eine Aussage. Für alle Änderungsoperationen gilt, daß sie Überzeugungszustände in Überzeugungszustände abbilden. Von solchen Zuständen fordern wir, daß sie deduktiv abgeschlossen seien. Was immer die Expansion $K + A$ sein soll, so ist jedenfalls klar, daß (1) $K + A = \mathrm{Cn}(K + A)$. Nun haben wir gesagt, daß eine Expansion von K um A eine Erweiterung von K sein soll, die A enthält. Damit fordern wir (2) $K \subseteq K + A$ und (3) $A \in K + A$. Was wir noch nicht erwähnt, aber sicher immer im Sinn haben, ist, daß eine Expansion um A nicht *irgendeine* Erweiterung von K sein soll, sondern eine, die nur den Zweck hat A zu K hinzuzunehmen. D.h., wenn A schon in K ist, dann gibt es für eine Expansion nichts zu tun: (4) Wenn $A \in K$, dann $K + A = K$. Ferner, wenn $K \subseteq K'$ und $K + A$ größer als $K' + A$ ist, dann würde $K + A$ über den Zweck hinausgehen; deshalb (5): Wenn $K \subseteq K'$, dann $K + A \subseteq K' + A$. Und schließlich sollte $K + A$ die genannten Bedingungen auf die sparsamste Weise erfüllen, d.h. (6) $K + A$ sei die kleinste Menge, die (1–5) erfüllt.

Wenn wir alle diese Bedingungen zusammen betrachten, dann gibt es einen offensichtlichen Kandidaten für die Definition von $K + A$:

$$K + A = \mathrm{Cn}(K \cup \{A\}). \tag{+}$$

Tatsächlich ist (+) nicht nur eine Möglichkeit unter anderen, $K + A$ so zu definieren, daß die Bedingungen (1–6) erfüllt sind. Vielmehr gilt die stärkere, erste Repräsentationsbeobachtung:

Satz 1. $K+A = \mathrm{Cn}(K\cup\{A\})$ *genau dann, wenn* $K+A$ *die Bedingungen* (1-6) *erfüllt.*

Beweis. LR: Sei $K + A = \mathrm{Cn}(K \cup \{A\})$. Dann ist leicht nachzuprüfen, daß $K + A$ die Bedingungen (1–5) erfüllt. Für (6) nehmen wir an – für *reductio* –, es gäbe eine ebenfalls (1–5) erfüllende Menge H mit $H \subset \mathrm{Cn}(K\cup \{A\})$. Dann gibt es ein B mit $K, A \vdash B$ und (nach 1) $H \nvdash B$. Also $K \vdash A \to B$. Aber $K \subseteq H$ (nach 2). Also $H \vdash A \to B$, und da $H \vdash A$ (nach 3), so folgt $H \vdash B$ – Widerspruch.

RL: Wir nehmen an, $K + A$ erfülle die Bedingungen. Zuerst zeigen wir, daß (∗) $\mathrm{Cn}(K \cup \{A\}) \subseteq K + A$. Nach (2) haben wir $K \subseteq K + A$, nach (3) $A \in K+A$. Also $K\cup\{A\} \subseteq K+A$. Es folgt nach Monotonie $\mathrm{Cn}(K\cup\{A\}) \subseteq \mathrm{Cn}(K + A)$ [$= K + A$ nach (1)].

Nun nutzen wir die Annahme, daß $K + A$ die Bedingung (6) erfüllt. Da wir in (LR) gezeigt haben, daß $\mathrm{Cn}(K \cup \{A\})$ eine Menge ist, welche die Bedingungen (1–5) erfüllt, so ist nach (6) ausgeschlossen, daß $\mathrm{Cn}(K \cup \{A\}) \subset K + A$. Es folgt aus (∗), daß $\mathrm{Cn}(K \cup \{A\}) = K + A$. ■

Das war einfach: Expansionen können rein logisch bestimmt werden. Wir können (+) als eine Definition verstehen und von sich daraus ergebenden Eigenschaften, wie (1–6), je nach Bedarf Gebrauch machen. Wir können aber auch mit den Expansionspostulaten (1–6) beginnen und dann beobachten, daß jede Operation + mit diesen Eigenschaften sich wie in (+) beschrieben, darstellen läßt. Nach diesem zweiten Muster werden wir nun Kontraktionen und Revisionen behandeln.

Kontraktion und Revision. Wir kommen nun zur Operation der *Kontraktion*. Auch Kontraktionen sollen Abbildungen in deduktiv abgeschlossene Mengen sein, d.h. wir werden fordern, daß $K - A = \mathrm{Cn}(K - A)$. Der Zweck einer Kontraktion um A ist es, die Überzeugung A aufzugeben. Also sollte $A \notin K - A$ der Fall sein. Hier ist jedoch zu beachten, daß durch den deduktiven Abschluß logische Wahrheiten in jedem Überzeugungszustand erzwungen sind. Also kann $A \notin K - A$ nur dann der Fall sein, wenn A keine logische Wahrheit ist. Kontraktionen nehmen etwas zurück, was zuvor typischerweise vorhanden war; also $K - A \subseteq K$. Im Falle, daß A in K gar nicht vorhanden ist, gibt es nichts, was die Kontraktion um A entfernen müßte; also in diesem Fall $K - A = K$. Wie schon Expansionen, so sollen auch Kontraktionen ihren Zweck auf möglichst sparsame Weise erzielen. Es sollen nicht mehr Überzeugungen als für den Zweck nötig entfernt werden. Dazu könnte folgender Test dienen: Wenn wir zu $K - A$ die Überzeugung A per Expansion wieder hinzunehmen, dann sollte der alte Zustand wiederhergestellt sein: $(K - A) + A = K$. Das kann natürlich nur im typischen Fall $A \in K$ gelten; denn anderenfalls wäre $K - A = K$ und so würde unsere Bedingung K mit $K + A$ gleichgesetzt. Schließlich soll es nicht darauf ankommen, wie die zu entfernende Überzeugung ausgedrückt ist. Logisch gleichwertige Sätze drücken dieselbe Überzeugung aus und bringen daher gleiche Kontraktionen hervor.

Was *Revisionen* angeht, so könnten wir uns einfach an Levis Idee halten (vgl. p. 488), daß jede legitime Revision aus einer Kontraktion, gefolgt von einer Expansion besteht, und zwar so:

Levi-Identität (LI) $$K * A = (K - \neg A) + A.$$

Wir wollen uns aber zunächst der Eigenschaften von Revisionen vergewissern, ohne auf die von Kontraktionen zurückzugreifen. Danach können wir prüfen, ob die beiden Gruppen von Eigenschaften die (LI) tatsächlich stützen.

Wie schon zuvor, wird $K * A = \mathrm{Cn}(K * A)$ zu fordern sein. Ferner sollten Revisionen auf zweierlei Weise erfolgreich sein: Erstens, sollten sie die Überzeugung um die revidiert wird, enthalten, und, zweitens, sollten sie das, wenn möglich, auf konsistente Weise tun. Der zweite Punkt unterscheidet Revisionen von Expansionen, denn bei letzteren wird eine Überzeugung ja nur blind, d.h. ohne Rücksicht auf die Widerspruchsfreiheit des Resultats, hinzugefügt. Expansionen bilden also so etwas wie eine obere Schranke für Revisionen: $K * A \subseteq K + A$. Wenn A auf keine widersprechende Überzeugung in K trifft, dann fallen Revision und Expansion um A zusammen. Schließlich sollte es wieder nicht darauf ankommen, wie Überzeugungen ausgedrückt werden, solange die Ausdrücke logisch gleich sind.

Soweit haben wir nur über Bedingungen gesprochen, die sich auf beliebige Aussagen A beziehen, um die kontrahiert oder revidiert werden soll. Wenn

AGM-Postulate für Kontraktionen

(C1)	Abschluß	$\mathrm{Cn}(K - A) \subseteq K - A$	Closure
(C2)	Erfolg	$A \notin K - A,$ falls $\nvdash A$	Success
(C3)	Schranke	$K - A \subseteq K$	Inclusion
(C4)	Null	Wenn $A \notin K,$ dann $K \subseteq K - A$	Vacuity
(C5)	Zurück	$K \subseteq (K - A) + A$	Recovery
(C6)	Kongruenz	Wenn $A \equiv B,$ dann $K - A = K - B$	Extensionality

Zusatzpostulate

(C7) $K - A \cap K - B \subseteq K - (A \wedge B)$

(C8) Wenn $A \notin K - (A \wedge B),$ dann $K - (A \wedge B) \subseteq K - A$

AGM-Postulate für Revisionen

(R1)	Abschluß	$\mathrm{Cn}(K * A) \subseteq K * A$	Closure
(R2)	Erfolg	$A \in K * A$	Success
(R3)	Schranke	$K * A \subseteq K + A$	Inclusion
(R4)	Erhaltung	Wenn $\neg A \notin K,$ dann $K + A \subseteq K * A$	Preservation
(R5)	Konsistenz	Wenn $\perp \in K * A,$ dann $\vdash \neg A$	Consistency
(R6)	Kongruenz	Wenn $A \equiv B,$ dann $K * A = K * B$	Extensionality

Zusatzpostulate

(R7) $K * (A \wedge B) \subseteq (K * A) + B$

(R8) Wenn $\neg B \notin K * A,$ dann $(K * A) + B \subseteq K * (A \wedge B)$

Von Kontraktionen zu Revisionen und v.v.

(LI) Levi-Identität $K * A = (K - \neg A) + A$

(HI) Harper-Identität $K - A = K \cap (K * \neg A)$

nun aber die Zielaussage von der Form einer Konjunktion oder Disjunktion ist, dann ist $K*(A\wedge B)$ gewissermaßen eine simultane Revision um die zwei Konjunkte A und B, so wie $K - A \vee B$ aus K beide Disjunkte, A und B, entfernen muß. [5] So stellt sich die Frage, wie sich solche Änderungen um konjunktiv oder disjunktiv zusammengesetzte Aussagen zu den entsprechenden Änderungen um ihre Bestandteile verhalten. Welche Beziehungen zwischen $K-A$ und $K-B$ einerseits und $K-A\wedge B$ andererseits (und ebenso für Revisionen) sollten wir erwarten? Eine Antwort geben die C- bzw. R-Postulate 7 und 8.

Die Postulate (R7–8) scheinen die Basispostulate (R3–4) zu verallgemeinern. Das ist aber nur bedingt der Fall. So folgt (R4) aus (R8) (und R3) nur unter der Bedingung, daß wir einen sehr schwachen Fall von (R4) voraussetzen, nämlich,

$$(\mathrm{R4}^{\top}) \qquad K \subseteq K * \top, \text{ wenn } \bot \notin K.$$

Diese Bedingung fordert (im Verein mit dem R3-Fall $K * \top \subseteq K$), [6] daß die Revision einer konsistenten Theorie um eine logische Wahrheit $\top$ ohne Effekt bleibt. Die Bedingung ist gewissermaßen in die logische Grammatik von Revisionen eingebaut; denn da Revisionen immer auf Theorien operieren, können diesen logische Wahrheiten nicht echt hinzugefügt werden. Nur im Grenzfall K = FML kann $K \subseteq K * \top$ offensichtlich nicht gelten, denn auch in diesem Fall soll die Revision $K * \top$ konsistent sein. Wir haben hier die beiden Erhaltungs-Postulate (R4) und (R4$^{\top}$) herausgegriffen, weil sie im Zusammenhang mit dem Levi- und dem Ramsey-Test besonders in den Blick geraten werden.

Für die Konstruktionen von Kontraktionen und Revisionen spielen die Zusatzpostulate eine interessante Rolle, auf die wir später zurückkommen werden. An dieser Stelle sei nur eine bemerkenswerte Bedingung ("Ventilierung") festgehalten, die sich vor dem Hintergrund der anderen Postulate als äquivalent zu (C7–8) erweist (siehe [7, Satz 6.5]):

$$K-(A\wedge B)=\begin{cases} K-A \text{ oder}\\ K-B \text{ oder}\\ K-A\cap K-B.\end{cases}$$

[5] Es scheint so, als ob Änderungen mit Blick auf mehr als eine Zielaussage im endlichen Fall auf Änderungen um Disjunktionen bzw. Konjunktionen reduziert werden können. Aber das ist tatsächlich eine offene Frage, welche am besten in einer echten Erweiterung der AGM-Theorie, der Theorie multipler Änderungen behandelt wird; siehe [88].

[6] Diesen Fall müssen wir voraussetzen, wenn wir die allgemeine Bedingung (R3) aus (R7) ableiten möchten.

In den letzten Absätzen haben wir die AGM-Postulate beschrieben und dabei versucht, sie plausibel zu machen. [7] Die Postulate umschreiben auf intuitiv einleuchtende Weise, was wir von der Rücknahme bzw. Revision von Überzeugungen erwarten. Solche Plausibilitätsüberlegungen rechtfertigen die Wahl genau dieser Postulate aber nur zum Teil. Wir habe auch prozedurale Vorstellungen davon, wie Rücknahmen und Revisionen auszuführen sind, und diese Vorstellungen können ihrerseits auf verschiedene Weise beschrieben werden. Erst wenn diese Vorstellungen und Beschreibungen zu den Postulaten passen, erhalten diese die gewünschte Stabilität.

Wir kommen nun zurück auf die Frage, ob die Levi-Identität (LI) Revisionen im Sinne der R-Postulate erzeugt unter der Annahme, daß Kontraktionen die C-Postulate erfüllen. Bevor wir die Frage beantworten, wollen wir auch den umgekehrten Weg, von Revisionen zu Kontraktionen betrachten. Eine Revision von K um $\neg A$ erzeugt einen Nachfolgezustand, in dem A (im typischen Fall) nicht enthalten ist, und der gegenüber K um $\neg A$ und allem was daraus folgt, erweitert ist. Wenn wir diesen Erweiterungsteil aus $K * \neg A$ herausschneiden, dann bleibt nur der Effekt, daß A aus K entfernt wurde (um K für $\neg A$ zu öffnen) – mit anderen Worten, eine Kontraktion um A. Diese Idee liegt der sogenannten

Harper-Identität (HI) $\qquad K - A = K \cap (K * \neg A)$

zugrunde. [8]

Beobachtung 2. *Die Postulate für Kontraktionen und Revisionen* modulo (LI) *bzw.* (HI)*:*

1. *Aus den Kontraktionspostulaten zusammen mit* (LI) *folgen die Revisionspostulate. (NB: Zurück* (C5) *ist für die Ableitung nicht nötig!)*
2. *Aus den Revisionspostulaten zusammen mit* (HI) *folgen die Kontraktionspostulate.*

Beweis. Die nötigen Verifikationen für die Basispostulate sind einfach. Als Illustration sei hier für Teilsatz 2 (Zurück) $K \subseteq K - A + A$ hergeleitet. Nach Anwendung der (HI) ist zu zeigen, daß $K \subseteq (K \cap (K * \neg A)) + A$. Angenommen (1) $B \in K$. Da $B \in K + A$ gdw $\neg A \vee B \in K$, ist nun zu zeigen, daß (a) $\neg A \vee B \in K$ und (b) $\neg A \vee B \in K * \neg A$. (a) folgt unmittelbar aus (1), und (b) folgt aus Erfolg (R2).

[7] Allgemeine Bedingungen für Expansionen und Revisionen finden sich schon 1978 in Gärdenfors' Aufsatz [96]. Die klassische Liste der AGM-Postulate für Revisionen und Kontraktionen – siehe nächste Seite – erschien erstmals 1982 im Druck (Gärdenfors [97]).

[8] Die Gleichung wird auch manchmal "Gärdenfors-Identität" genannt. Wir halten uns hier an die Regel, möglichst die zuerst eingeführte Bezeichnung zu verwenden. Gärdenfors [96] verweist auf eine Idee Harpers in [131].

Die gegenseitige Herleitung der Zusatzpostulate ist etwas aufwendiger. Wir zeigen beispielhaft den Fall (C7) $\Rightarrow$ (R7) via (LI). — Angenommen $C \in K * (A \wedge B)$, d.h. (LI!)

$$A \wedge B \to C \in K - \neg(A \wedge B). \tag{1}$$

Zu zeigen: $(*)$ $A \wedge B \to C \in K - \neg A$. Um (C7) anzuwenden, ersetzen wir $\neg A$ in $(*)$ durch die äquivalente Formel $\neg(A \wedge B) \wedge (\neg A \vee B)$. Nach (C7) haben wir dann

$$K - \neg(A \wedge B) \cap K - (\neg A \vee B) \ \subseteq \ K - \neg(A \wedge B) \wedge (\neg A \vee B),$$

und es bleibt zu zeigen, daß

$$A \wedge B \to C \in K - \neg(A \wedge B) \quad \text{und} \quad A \wedge B \to C \in K - (\neg A \vee B).$$

Der linke Teil ist durch (1) gegeben. Nach Zurück (C5) haben wir

$$K \subseteq K - (\neg A \vee B) + (\neg A \vee B). \tag{2}$$

Da $A \wedge B \vdash \neg A \vee B$, so folgt aus (2),

$$K \subseteq K - (\neg A \vee B) + (A \wedge B). \tag{3}$$

Aus (1) und Schranke (C2) folgt, daß $A \wedge B \to C \in K$. Daher aus (3), $A \wedge B \to C \in K - (\neg A \vee B) + (A \wedge B)$, d.h. $A \wedge B \to (A \wedge B \to C) \in K - (\neg A \vee B)$, was äquivalent ist zum gewünschten $A \wedge B \to C \in K - (\neg A \vee B)$. ■

Die beiden Identitäten vermitteln also erfolgreich zwischen den beiden Gruppen von Postulaten. Wir können die Beobachtung auch im Sinne eines Repäsentationsresultates lesen:

1. Gegeben eine Operation $-$, welche die C-Postulate erfüllt, so erfüllt die Operation $*$, definiert nach (LI), die R-Postulate; und
2. Gegeben eine Operation $*$, welche die R-Postulate erfüllt, so erfüllt die Operation $-$, definiert nach (HI), die C-Postulate.

Angenommen, wir beginnen mit einer AGM-Kontraktionsoperation $-$. Dann erzeugen wir mit der (LI) eine Operation $\ell(-)$. Nun können wir auf der Basis von $\ell(-)$ nach (HI) eine Operation $h(\ell(-))$ definieren. Die Beobachtung garantiert, daß $h(\ell(-))$ eine AGM-Revisionsoperation ist.

Aber ist $h(\ell(-)) = -$? Und ebenso stellt sich die Frage, ob $\ell(h(*)) = *$ ist.

Wir fragen hier danach, ob die Menge $K - A$ identisch ist mit der Menge, die entsteht, wenn wir $K - A$ zunächst mit HI als $(K * \neg A) \cap K$ und dann die darin enthaltene Revision nach LI als $K - A + \neg A$ darstellen. Ähnlich für $K * A$ mit ungekehrter Anwendung der beiden Identitäten. D.h. wir fragen nach den Gleichungen

$$(h\ell) \qquad K - A = (K - A + \neg A) \cap K, \text{ und}$$

$$(\ell h) \qquad K * A = (K * A \cap K) + A.$$

Tatsächlich folgen $(h\ell)$ und (ℓh) aus den AGM-Postulaten.

Ad $(h\ell)$: Für die Inklusion von links nach rechts brauchen wir nur (C3). Für die umgekehrte Inklusion nehmen wir an (1) $B \in K$, und (2) $B \in K - A + \neg A$. Aus (1) folgt nach (C5), (3) $B \in K - A + A$. Aus (2) und (3) folgt, daß sowohl $A \wedge B$ als auch $\neg A \wedge B$ in $K - A$ sind. Also $B \in K - A$.

Ad (ℓh): Angenommen $B \in K * A$. Dann ist (R3!) B auch in $K + A$. Es folgt, daß $A \to B \in K * A \cap K$ und also $B \in (K * A \cap K) + A$. Für die umgekehrte Inklusion nehmen wir an, daß $A \to B \in K * A$. Dann folgt mit (R2) $B \in K * A$.

Kontraktionen und Revisionen sind also in einem starken Sinne wechselseitig definierbar. Wenn wir die Konstruktionsaufgabe für eine der beiden Operationen lösen, so lösen wir sie daher zugleich für die andere.

Das Zurück-Postulat (Recovery). Dieses Postulat ist das vergleichsweise problematischste in der AGM-Theorie. Im Falle $A \in K$, folgt aus Zurück und Schranke

$$K - A + A = K,$$

d.h. der ursprüngliche Zustand kann nach einer Kontraktion durch eine entsprechende Expansion wiederhergestellt werden.

Unter bestimmten Annahmen scheint diese Art der Wiederherstellung aber gar nicht möglich zu sein. Um A aus K zu entfernen, entfernen wir meist stärkere Elemente in K, die A implizieren. Wenn wir das schwächere A dann wieder hinzufügen, warum sollten wir dann die stärkeren Elemente zurückerhalten? Ein kleines formales Beispiel führt das deutlich vor Augen. Man betrachte die Überzeugungsmenge $K = \{P, Q \wedge R\}$. Als Kontraktion $K - Q$ scheint die Menge $\{P\}$ ein guter Kandidat zu sein. Wenn wir nun Q wieder hinzufügen, dann erhalten wir die Menge $\{P, Q\}$, welche deduktiv schwächer ist als K.

Das Gegenbeispiel zur Wiederherstellung funktioniert aber nur, wenn wir annehmen, daß die Kontraktion nicht auf einer deduktiv abgeschlossenen Theorie, sondern auf einer nicht deduktiv abgeschlossenen, meist endlichen Menge durchgeführt wird. Es spricht einiges dafür, solche Kontraktionen zu betrachten, denn tatsächlich bedienen wir uns ja bei der Arbeit mit Theorien immer einer endlichen Basis, sei es in Form einer Menge von Modellen oder einer Menge sogenannter Axiome und Regeln – beide zumeist übersichtlich beschreibbar. Es ist daher in der Anwendung naheliegend, Theorienänderungen als vermittelt über direkte Änderungen einer solchen axiomatischen Basis aufzufassen. Wenn wir das tun, dann müssen wir Kontraktionen und Revisionen für beliebige – jedenfalls nicht nur für deduktiv abgeschlossene – Mengen beschreiben. In diesem Fall können wir die genannte Eigenschaft der Wiederherstellung, und also Zurück nicht erwarten.

Eine Verallgemeinerung der AGM-Theorie, die Änderungsoperationen für beliebige Mengen definiert, steht im wesentlichen vor zwei Herausforderungen. Die erste – und relativ einfach zu lösende – besteht darin, Ersatz für Zurück zu finden. Wir erinnern uns, daß Zurück die wichtige Bedingung formuliert, daß eine Kontraktion nur minimal in die Menge der Überzeugungen eingreifen soll: gerade soviel, wie nötig ist, um eine bestimmte Überzeugung zurückzuziehen. Die zweite Herausforderung ist prinzipiellerer Natur. Eine Theorie kann auf unendlich viele Weisen axiomatisiert werden. Wenn es uns darum geht, eine Aussage aus einer Theorie zurückzunehmen, dafür aber eine bestimmte Basis als Mittel ausgezeichnet werden muß, dann hängt das Resultat vielleicht zu sehr von der Gestalt dieser Basis ab. So macht es in unserem kleinen Beispiel einen großen Unterschied, ob die Menge K mit $\{P, Q \wedge R\}$ oder mit $\{P, Q, R\}$ axiomatisiert ist. In den meisten Fällen würden wir jedoch sagen, daß zwischen der Zuschreibung der einen Überzeugung $P \wedge Q$ und der zweier Überzeugungen P und Q kein Unterschied besteht. Theorienänderungen mittels Basisänderungen sind in einer Weise syntax-empfindlich, wie es nicht paßt zu unserer Einstellung, daß die Wahl der Darstellung einer Theorie meist nur eine Frage der Bequemlichkeit ist. Trotz dieser prinzipiellen Schwierigkeit ist die Theorie der Basisänderungen eine bedeutende Erweiterung der AGM-Theorie. Sie ist jedoch nicht Gegenstand dieses Kapitels und der Leser sei auf die Arbeiten [84, 85, 125, 126] verwiesen.

In der AGM-Theorie nehmen wir an, daß es für die angestrebte Änderung einer Theorie keine Rolle spielen sollte, wie diese axiomatisiert ist. Wenn die AGM-Operationen sich daher auf logisch abgeschlossene Theorien richten, dann ist Zurück nicht nur sehr plausibel, sondern beinahe zwingend. Denn für jede Überzeugung B in K haben wir durch deduktiven Abschluß auch $\neg A \vee B \in K$. Wenn es nur darum geht, A aus K zu entfernen, dann würde es über diesen Zweck hinausgehen, die schwächere Überzeugung $\neg A \vee B$ zu löschen; d.h. $\neg A \vee B$ sollte in $K - A$ verbleiben. Aber wenn wir dann wieder A hinzufügen, erhalten wir durch logischen Abschluß auch B zurück:[9]

$$\dfrac{\dfrac{\begin{array}{c}\dfrac{B \in K}{\neg A \vee B \in K}\\ \vdots\ \text{minimaler Einschnitt!}\\ \neg A \vee B \in K - A\end{array}}{\neg A \vee B \in K - A + A} \qquad\qquad A \in K - A + A}{B \in K - A + A}$$

Zurück drückt die Bedingung aus, daß wir so argumentieren können. Neben Null ist Zurück das einzige Postulat, das der Maxime Geltung verschafft, daß Kontraktionen Information nicht ohne Not preisgeben und in diesem Sinne nur minimale Einschnitte in Theorien vornehmen sollten.[10]

Um uns eine Vorstellung davon zu machen, was ohne Zurück möglich wäre, betrachten wir einmal Operationen, die alle C-Postulate, jedoch nicht (C5), Zurück, erfüllen; Makinson [195] nennt sie *Entzugsoperationen* (*withdrawal operations*). Hier ist eine solche Entzugsoperation:

$$(\dagger) \qquad K - A = \mathrm{Cn}(\emptyset).$$

Diese Operation läßt (beinahe) nichts von K übrig! Sie verletzt maximal die Maxime des minimalen Eingriffs. Wir brauchen also unter den AGM-Postulaten eine Bedingung, welche solche Kontraktionen ausschließt. Makinson [195] zeigt, daß Zurück diejenige Bedingung ist, welche die genannte Maxime in bester, d.h. Information maximierender Weise umsetzt.

Aber nun könnte man fragen, ob eine wesentliche Annahme des Argumentes für Zurück zutrifft: Ist das Maximieren von Information im Sinne eines Inklusionsvergleichs von Aussagenmengen überhaupt das richtige Ziel? Könnte es nicht sein, daß zwei Theorien, H und H' mit $H \subset H'$, gleichwertig sind im Hinblick auf einen *qualitativen* Informationsvergleich? In diesem Fall würde H' gegenüber H einen Überschuß wertloser Information enthalten.

[9] Zurück ist übrigens das einzige Prinzip, zu dessen Verifizierung die AGM-Theorie von logischen Annahmen Gebrauch machen muß, die z.B. in der Relevanzlogik nicht gelten: Wir schließen gleich nach dem Disjunktiven Syllogismus von $\neg A \vee B$ und A auf B. Vgl. auch den Beweis von Satz 9, wo wir $A \to (\neg A \to B)$ annehmen müssen, um Zurück zu verifizieren.

[10] In der Beweisskizze oben haben wir das Prinzip minimaler Einschnitte vorausgesetzt. Im Beweis von Satz 9 wird das Prinzip für eine bestimmte Art, Kontraktionen zu konstruieren, bewiesen: Es taucht dort als Beobachtung $(*)$ auf.

Genau diese Möglichkeit möchte Levi [174, 176] offen halten und kommt so zu dem Schluß, daß Zurück keine notwendige Bedingung für gute Kontraktionen sein kann.[11] Auch die Harper-Identität (HI) muß nun fallen. Denn aus Erfolg (R2) und (HI) folgt Zurück (Beobachtung 2.2). So entsteht eine Asymmetrie zwischen Kontraktionen und Revisionen: Revisionen lassen sich über Kontraktionen definieren aber nicht umgekehrt – ein Resultat, das zu Levis Überlegungen über gerechtfertigte Überzeugungsänderungen gut paßt.

Die Präsenz von Zurück in der AGM-Theorie weist darauf hin, daß sie die Maxime der minimalen Einschnitte in einfacher Weise umsetzt, d.h. im Sinne eines Vergleichs des logischen Gehalts von Theorien und ohne den Qualitätsparameter, den Levi ins Spiel bringen möchte. Wie immer man zu Zurück stehen mag, jedenfalls spielt die Bedingung keine Rolle bei der Herleitung der Revisionspostulate über die Levi-Identität; siehe das *NB* zur Beobachtung 2.1. Selbst die in (†) definierte Entzugsfunktion generiert über (LI) eine Revisionsfunktion, die alle AGM-Postulate erfüllt. Aber in diesem Fall ist für alle K und A, $K*A = \mathrm{Cn}(A)$ – d.h. das Resultat fällt viel zu klein aus! Das Optimum – im leider kontroversen Sinne – bieten hier genau solche Entzugsfunktionen, die Zurück erfüllen, also die AGM-Kontraktionen.

5. Der Levi- und der Ramsey-Test

In diesem Abschnitt wollen wir auf einige Grenzen der Theorie eingehen. An diese Grenzen stoßen wir, wenn wir die Sprache, in der Überzeugungen formuliert werden, um Junktoren erweitern, die selbst auf Theorien oder deren Änderungen Bezug nehmen. Wenn eine Theorie K die Aussage A nicht enthält, dann hat K die Eigenschaft, mit $\neg A$ verträglich zu sein. Diese Eigenschaft könnte selbst Gegenstand einer Überzeugung sein. Es wäre eine Überzeugung, zu der man durch bloße Reflexion über K gelangen kann. Sollte K also die Aussage enthalten, daß $\neg A$ möglich ist? Oder, wenn B in $K * A$ ist, dann hat K die Eigenschaft, nach Revision um A die Aussage B zu enthalten. Auch das ist eine Eigenschaft, von der man *a priori* Kenntnis haben kann. Ist K also zur Aussage 'Wenn A der Fall wäre, dann wäre auch B der Fall' verpflichtet? Wie sich herausstellen wird, ist äußerste Vorsicht geboten.

---

[11] Siehe hierzu auch die Arbeiten von Pagnucco und Rott [253, 251].

Der Levi-Test für reflektive Modalitäten. Ein Überzeugungssystem K gibt einen Standard ab, nach dem beurteilt werden kann, was eine ernsthafte Möglichkeit ist und was nicht. Wenn $A \in K$, dann scheidet, aus der Sicht von K, die Aussage $\neg A$ aus dem Kreis ernsthafter Möglichkeiten aus. Und umgekehrt ist es so, daß wenn A keine ernsthafte Möglichkeit für K ist, dann liegt das daran, daß K schon auf $\neg A$ festgelegt ist. Schreiben wir $\Diamond A$ für "A ist eine ernsthafte Möglichkeit", dann könnten wir versucht sein, das gerade Gesagte so zusammenfassen:[12]

Levi-Test (LT): $\Diamond A \in K$ gdw $\neg A \notin K$.

Wir wollen einen Überzeugungszustand K *vollständig* nennen, wenn für jede Aussage A entweder A oder $\neg A$ in K ist. Wer sich in einem vollständigen Überzeugungszustand befindet, hat für sich also jede Ja/Nein-Frage entschieden. Natürlich kommen vollständige Überzeugungszustände praktisch nicht vor. Auf den ersten Blick scheint der (LT) die Möglichkeit unvollständiger Überzeugungszustände zuzulassen. Wenn weder A noch $\neg A$ in K sind, dann fordert der (LT) dies in K so zu registrieren, daß sowohl A und $\neg A$ ernsthafte Möglichkeiten darstellen, also $\Diamond A$ und $\Diamond\neg A$ in K sind. Wenn wir den (LT) jedoch in die AGM-Postulate einbetten, dann erweisen sich unvollständige Überzeugungszustände als unmöglich.

Beobachtung 3. *Der Levi-Test* (LT), *Erfolg* (R2) *und Erhaltung* (R4) *implizieren Vollständigkeit.*

Beweis. Für *reductio* nehmen wir an, K sei unvollständig bezüglich einer Aussage A und leiten dann so einen Widerspruch ab:

$$\frac{\dfrac{}{A \in K * A}\,\text{R2}}{\Diamond\neg A \notin K * A}\,\text{LT} \qquad \frac{\dfrac{A \notin K}{\Diamond\neg A \in K\ [\subseteq K + A]}\,\text{LT} \quad \dfrac{\neg A \notin K}{K + A \subseteq K * A}\,\text{R4}}{\Diamond\neg A \in K * A} \quad \blacksquare$$

Der Ramsey-Test für kontrafaktische Konditionale. Die Revision eines Überzeugungssystems um K, also $K * A$, gibt einen Standard ab, nach dem beurteilt werden kann, zu welchen Konditionalen K verpflichtet ist. Wenn $B \in K * A$, dann sollte aus der Sicht von K das Konditional $A \sqsupset B$ ("Wäre A der Fall, würde auch B der Fall sein.") gelten. Und umgekehrt, ist $A \sqsupset B$ aus der Sicht von K richtig, dann sollte B in $K * A$ sein. Wie

[12] Man beachte, daß Levi ausdrücklich vor diesem Test warnt! Der Test wird hier jedoch so benannt, weil der Begriff der ernsthaften Möglichkeit, im gerade erklärten Sinne, in Levis Theorie eine zentrale Rolle spielt. Levi zufolge ist die bessere Version des Tests die Äquivalenz (LLT), welche wir weiter unten vorstellen werden.

schon einmal erklärt (489ff), liegt es nahe, diesem Gedanken die folgende Form zu geben:

Ramsey-Test (RT): $A \sqsupset B \in K$ gdw $B \in K * A$.

Wie (LT), so kann auch (RT) nicht die AGM-Postulate ergänzen. Aber noch bevor wir bestimmte Postulate im Verein mit (RT) betrachten, können wir schon sehen, daß Revisionen nicht unter die Bedingung (RT) gestellt werden sollten. Denn die Bedingung impliziert unmittelbar, daß Revisionen monoton sind:

Monotonie (RM): Wenn $K \subseteq K'$, dann $K * A \subseteq K' * A$.

Warum ist das unplausibel? Nun, Revisionen unterscheiden sich wie Kontraktionen von Expansionen dadurch, daß sie ein Element der Wahl enthalten. Um zur Revision $K * A$ zu gelangen, müssen wir entscheiden, welche Aussagen in K wir aufgeben wollen, um unsere Überzeugungen für A zu öffnen. Da eine Obermenge H von K mehr Information enthält, könnten wir im Lichte dieser Information die aus K zu entfernenden Aussagen anders wählen, als wir es täten, wenn wir nur auf die Information in K beschränkt sind. Aber genau das schließt (RM) aus.

Obgleich das schon ein gutes Argument gegen (RM) ist, treibt das nächste gewissermaßen den letzten Nagel in den Sarg. Es greift auf elementare Postulate für Revisionen, nämlich auf Erhaltung (R4) und Konsistenz (R5) zurück.

Wenn in (R4) und (R5) A die *Verum*-Konstante $\top$ ist, dann besagt (R5) einfach, daß $K * \top$ konsistent ist, und (R4) besagt, daß $K \subseteq K * A$, falls K konsistent ist. Nun können wir folgendes zeigen:

- Unter den Bedingungen (RM) und (R4–5) kann es keine Formeln geben, die für sich konsistent, aber zusammen inkonsistent sind.

Seien A und A' solche Formeln, d.h. (1) $A \nvdash \bot$ und (2) $A' \nvdash \bot$, jedoch (3) $A, A' \vdash \bot$. Wir können folgende Theorien bilden:

$$K = \mathrm{Cn}(A), \quad K' = \mathrm{Cn}(A'), \text{ und } \quad H = \mathrm{Cn}(A, A').$$

Es wird genügen, zu zeigen, daß sowohl A als auch A' in $H * \top$ sind. Denn dann ist nach (3) $H * \top$ inkonsistent – im Widerspruch zu (R5). — Da K nach Annahme (1) konsistent ist, haben wir $K \subseteq K * \top$ nach (R4) und also ist A in $K * \top$. Da $K \subseteq H$, so folgt nach (RM), $K * \top \subseteq H * \top$ und also $A \in H * \top$. Genau analog erhalten wir mit Annahme (2) auch $A' \in H * \top$.

Das Trivialisierungsresultat von Gärdenfors ist etwas allgemeiner, da es nicht darauf ankommt, ob wir in (1–3) die Konstante $\bot$ oder eine beliebige Aussage B wählen.

Beobachtung 4. (Gärdenfors [99], Segerberg [262].) *Unter den Bedingungen Monotonie* (RM)*, Erhaltung* (R4) *und Konsistenz* (R5) *kann es keine Sätze A, A' und B geben so, daß* (1) $A \nvdash \neg B$, (2) $A' \nvdash \neg B$, *und* (3) $A, A' \vdash \neg B$.

Beweis. Das Argument ist im wesentlichen so, wie gerade vorgestellt. K, K' und H sind wie oben definiert. Dann argumentieren wir so:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \in K}{A \in K + B} \qquad B \in K + B}{A, B \in K + B} \qquad \dfrac{(1)}{K + B \subseteq K * B}\,\mathrm{R4}}{A, B \in K * B} \qquad \dfrac{K \subseteq H}{K * B \subseteq H * B}\,\mathrm{RM}}{A, B \in H * B}$$

Auf dieselbe Weise (mit Annahme (2) statt (1), A' statt A, und K' statt K) kommen wir zur Konklusion $A', B \in H * B$. So haben wir $A, A', B \in H * B$. Aber dann folgt aus (3), daß $\perp \in H * B$. Aber aus (1) (und auch aus 2) folgt, daß $\nvdash \neg B$; also nach (R5), $\perp \notin H * B$—Widerspruch! ■

In nicht-trivialen Sprachen kann also der Ramsey-Test die AGM-Postulate nicht ergänzen. Dennoch bleibt die Erzeugung einer genügend starken Konditionallogik auf der Basis des Tests unbeschadet, denn dafür können wir Erhaltung (R4) – und auch die bedingte Verallgemeinerung (R8) – umgehen. Für den folgenden Satz wollen wir unter einem *Modell* ein Paar $(\mathcal{K}, *)$ verstehen. Darin ist $\mathcal{K}$ eine Menge von Theorien, die unter Expansion und Revision $*$ abgeschlossen ist. Im allgemeinsten Sinn eines Modells nehmen wir nur an, daß $*$ eine Abbildung von $\mathcal{K} \times \mathrm{FML}$ nach $\mathcal{K}$ ist so, daß Bedingung (R6) erfüllt ist, d.h. [13]

$$\text{wenn } A \equiv B, \text{ dann } K * A = K * B.$$

Die Theorien eines Modells sind Teilmengen von Formeln in einer Boole'schen aussagenlogischen Sprache, die mit einem zweistelligen Junktor $\sqsupset$ angereichert ist. Daraus ergeben sich die Bedingungen, unter denen eine Theorie wahrheitsfunktional zusammengesetzte Formeln enthalten kann. Für $\sqsupset$ ist (RT) zuständig. Man beachte, daß Theorien unvollständig sein können, d.h. aus $\neg A \notin K$ folgt nicht, daß $A \in K$. Deshalb definieren wir gleich Gültigkeit im Sinne von Unwiderlegbarkeit. Ferner gibt es in jedem Modell genau eine Theorie, die inkonsistent ist.

[13] Das ist etwas weniger allgemein als in Gärdenfors [96, 100]. Aber die Bedingung ist die naheliegendste überhaupt und erlaubt es, umstandslos an unsere Behandlung der Konditionallogik in Kapitel IV anzuschließen.

- Eine Formel A ist *unwiderlegt in einem Modell* $(\mathcal{K}, *)$, wenn es kein konsistentes $K \in \mathcal{K}$ gibt, mit $\neg A \in K$.
- Eine Formel ist genau dann *gültig in einer Modellklasse*, wenn sie darin unwiderlegbar, d.h. in allen Modellen dieser Klasse unwiderlegt ist.

SATZ 5. (Gärdenfors [96].) *Die Menge der in der Klasse aller Modelle gültigen Formeln ist die Menge der Theoreme der Logik* **CK** *(siehe Kap. VI.3).*

Indem wir nun die Operation $*$ unter weitere Bedingungen stellen, können wir weitere Konditionallogiken erzeugen. Erhaltung (R4) sollten wir dazu natürlich nicht wählen. Stattdessen dürfen wir feststellen, daß die schwächere und unschädliche Bedingung

(R4′) Wenn $A \in K$ und $\perp \notin K$, dann $K \subseteq K * A$

(aus der übrigens (R4$^\top$) folgt – siehe die Diskussion p. 496) genau mit dem sogenannten Zentrierungsschema

CS. $$A \wedge B \rightarrow (A \sqsupset B)$$

korrespondiert. Auch für das Schema

CV. $$(A \sqsupset B) \wedge (A > C) \rightarrow (A \wedge C \sqsupset B)$$

genügt eine schwächere Version von (R8), nämlich

(R8′) Wenn $\neg(A \sqsupset \neg B) \in K$, dann $K * A + B \subseteq K * (A \wedge B)$.

So können wir alle im Kapitel VI beschriebenen Konditionallogiken sukzessive aufbauen, bis hin zum starken Lewis-System **VC**. Wenn wir in den R-Postulaten (R4) und (R8) durch die Abschwächungen (R4′) und (R8′) ersetzen, dann legt diese schwächere Gruppe von Bedingungen eine Klasse AGM′ von Modellen festlegt.

SATZ 6. (Gärdenfors [96].) *Die Menge der in der Klasse* AGM′ *gültigen Formeln ist die Menge der Theoreme der Logik* **VC**.

Wir können über **VC** noch hinausgehen zur Stalnaker-Logik **VCS**, welche **VC** das Schema

CEM. $$(A \sqsupset B) \vee (A \sqsupset \neg B)$$

hinzufügt. Dieses Schema spiegelt eine Bedingung, die aus den AGM-Postulaten *nicht* ableitbar ist:

$K * A$ ist vollständig, wenn K vollständig ist.

(Eine Theorie ist vollständig, falls sie für jede Formel A entweder A oder $\neg A$ enthält.)

* * *

Die beiden Trivialisierungsresultate haben offenbar etwas gemein. Schon bei oberflächlicher Betrachtung sehen wir, daß Erhaltung (R4) die einzige Bedingung ist, die in beiden Ableitungen wesentlich vorkommt. Ferner definieren die Tests für $\diamond$ und $\sqsupset$ Junktoren, welche auf systematische, d.h. richtige und vollständige Weise Feststellungen *über* Überzeugungszustände in Überzeugungen selbst spiegeln; die beiden Junktoren stellen also so etwas wie "reflektive Modalitäten" dar [82]. Die Erhaltung-Bedingung, die in der Abwesenheit reflektiver Modalitäten keinen Anlaß zur Beanstandung gibt, führt im Verein mit (LT) oder (RT) zu einem Desaster. Ein offensichtlicher Ausweg könnte daher darin bestehen, das Postulat (R4) (und auch R8) anzupassen. Die Abschwächung zu (R4′) (und R8′) ist kaum mehr als eine technische Lösung, die sich für die Erzeugung von Konditionallogiken eignet.[14]

Für eine alternative Lösung spricht allein schon der Umstand, daß der (RT) unmittelbar, d.h. ohne (R4), die an sich unplausible Monotonie-Bedingung (RM) impliziert. Levi setzt daher bei den Tests an. Der Umstand, daß ein Korpus K eine Aussage A als ernsthafte Möglichkeit offen hält bzw. K zur Aufnahme von B in $K * A$ verpflichtet ist, ist zunächst einmal nur eine Eigenschaft von K. Wenn uns daran liegt, können wir Aussagen über solche "modalen" Eigenschaften von K in einem Metakorpus $M(K)$ nach diesen Definitionen festhalten:

LLT $\qquad \diamond A \in M(K) \text{ gdw } \neg A \notin K;$

LRT $\qquad A \sqsupset B \in M(K) \text{ gdw } B \in K * A.$

Die oben angestellten Beobachtungen warnen uns jedoch davor, die Ebenen von Objekt- und Metakorpus zu vermengen. Für Levi (siehe insbesondere [175]) gibt es einen weiteren, prinzipielleren Grund, warum ein solcher Kollaps keinem guten Zweck dienen kann. Ein Korpus K von Hypothesen soll Ausgangspunkt sein für Änderungen unter der Vorgabe Irrtum zu meiden. Aber keine echte Verkleinerung und keine echte Erweiterung von K kann diese Vorgabe erfüllen. Denn jede Verkleinerung $K^- \subset K$ öffnet eine Möglichkeit $\diamond A$, die in K keine gewesen ist ($\neg\diamond A \in K$). Wenn K^- unter (LT) abgeschlossen ist, dann ist $\diamond A \in K^- \subseteq K$ – was nur möglich ist, wenn K inkonsistent ist. Der (LT) würde also verhindern, daß konsistente Überzeugungszustände kontrahiert werden können. Ebenso verschließt jede Erweiterung $K^+ \supset K$ eine Möglichkeit $\diamond A$ (also $\neg\diamond A \in K^+$), welche in K noch gegeben ist ($\diamond A \in K$) – was nur möglich ist, wenn K^+ inkonsistent

[14] Eine grundsätzliche Diskussion der Erhaltung-Bedingung ist in [82], [249] und [109] zu finden.

ist. Der (LT) würde also verhindern, daß ein Überzeugungszustand jemals konsistent erweitert werden könnte. Ähnlich argumentiert Levi gegen den Abschluß unter (RT).

Beide Arten modaler Aussagen, $\Diamond A$ and $A \sqsupset B$, haben ihre Funktion im Prozess der Überzeugungsänderung: $\Diamond$-Aussagen legen den Spielraum induktiver Erweiterungen, $\sqsupset$-Aussagen den legitimer Revisionen fest. Aber sie können nicht Teil eines Korpus K sein, das als Ausgangspunkt legitimer Überzeugungsänderungen dienen soll – jedenfalls dann nicht, wenn ihre Bedeutungen durch die Test-Äquivalenzen festgelegt sind.

Ein Ansatz, der formal gesehen kaum mehr als eine Notationsvariante des Ansatzes mit Metakorpora ist, wird in [64] vorgestellt: Den Bezug der modalen Aussagen auf ein gegebenes Korpus K können wir auch durch indizierte Junktoren deutlich machen und diese so einführen:

$$\Diamond_K A \in K \text{ gdw } \neg A \notin K.$$

Durch die Indizierung läßt sich der Widerspruch im Beweis von Beobachtung 3 nicht mehr herleiten. Ebenso können wir mit Konditionalen verfahren (wie schon in [187] ausgeführt) und bleiben dann bereits in der Herleitung der Monotonie von Revisionen stecken.

6. Konstruktionen mit Teilschnitten (partial meet)

Nachdem wir bisher Bedingungen betrachtet haben, die Kontraktionen und Revisionen erfüllen sollten, wenden wir uns nun der anderen Seite der AGM-Theorie zu: der Angabe von Lösungswegen (Konstruktionen) für Kontraktions- bzw. Revisionsaufgaben. Für die Entwicklung der AGM-Theorie spielen die Konstruktionen mit Teilschnitten (*partial meet constructions*) eine zentrale Rolle; sie werden in diesem Abschnitt behandelt.

Grundlegende Definitionen. Kontraktionen, so haben wir gesagt, sollen nur minimale Einschnitte in Theorien vornehmen: gerade so viel, wie nötig ist, um den Zweck der Kontraktion – die Aufgabe einer bestimmten Überzeugung – zu erfüllen. Wenn wir eine Kontraktionsaufgabe $K - A$ lösen wollen, besteht daher der erste Schritt darin, die *Kandidaten*, welche diese Bedingung in bester Weise erfüllen, zu bestimmen.

1. Bestimme alle maximalen Kandidaten für $K - A$, d.h. alle maximalen Teilmengen von K, aus denen A nicht folgt.

Im nächsten Schritt ist ein Wahlproblem zu lösen:

2. Wähle aus: (A) alle Kandidaten, (B) den besten Kandidaten, oder (C) alle besten Kandidaten.

Die Alternativen (B) und (C) setzen voraus, daß wir aus den Kandidaten eine *Auswahl* treffen können. Es ist daher naheliegend, daß wir uns die Kandidaten im Sinne einer durch das Überzeugungssystem K vorgegebenen Präferenzordnung angeordnet denken. Alternative (B) setzt darüber hinaus voraus, daß diese Ordnung für jede Überzeugung A immer genau einen besten Kandidaten für K ohne A auszeichnet. Das ist eine starke Annahme, welche nicht nur für sich betrachtet wenig pausibel ist, sondern – wie wir gleich sehen werden – kontrahierte Überzeugungsmengen generell viel zu klein macht.

Die Alternativen (A) und (C) ergeben noch keinen bestimmten Nachfolger $K - A$ von K. Um die Kontraktionsaufgabe ohne Willkür zu lösen, sollten wir danach fragen, worin alle Kandidaten übereinstimmen. Das ist die sicherste Art und Weise, die Aussage A zurückzuziehen, d.h. $K - A$ ist der *Schnitt* aller infrage kommenden Kandidaten.

Die folgende Definition faßt diese Schlüsselideen zusammen. Aus den Mengen, die wir gerade "Kandidaten" genannt haben, wählt eine Funktion s (abhängig von K) eine Teilmenge aus, welche dann in sich geschnitten wird.

Definition 7. Es sei K eine Überzeugungsmenge.

1. Die *Restefamilie* (*remainder set*) $K \perp A$ für K ohne A besteht aus allen Mengen X so, daß
   (a) $X \subseteq K$;
   (b) $X \nvdash A$; und
   (c) $\forall Y$: wenn $X \subset Y \subseteq K$, dann $Y \vdash A$.
2. Es sei s eine *Auswahlfunktion* (für K) so, daß für jede Aussage A,
   (a) $s(K \perp A) \subseteq K \perp A$; und
   (b) $s(K \perp A) = \begin{cases} \neq \emptyset, & \text{falls } K \perp A \neq \emptyset; \\ s(K \perp A) = \{K\}, & \text{anderenfalls.} \end{cases}$
3. $K - A = \bigcap s(K \perp A)$ ist die Kontraktion K ohne A nach *Teilschnitt* (*partial meet*).

Teil 3 der Definition gibt die Option (C) oben wieder. Die beiden anderen Optionen, (A) und (B), ergeben sich aus (C) als Spezialfälle, indem wir jeweils die Auswahl s (für eine gegebene Theorie K) unter entsprechende Bedingungen stellen:

4. Wenn $s(K \perp A) = K \perp A$, dann nennen wir $K - A$ eine Kontraktion nach *vollem Schnitt* (*full meet*). (In diesem Fall ist die Auswahlfunktion s redundant.)
5. Wenn aus für alle $X, X' \in s(K \perp A)$ folgt, daß $X = X'$, dann nennen wir $K - A$ eine Kontraktion nach *Maximalwahl* (*maxichoice*). (In

diesem Fall wählt die Auswahlfunktion s aus jeder Restefamilie genau ein Element.)

Voller Schnitt und Maximalwahl erweisen sich als keine guten Kontraktionsstrategien. Zwar erfüllen sie die AGM-Postulate, haben darüber hinaus aber Eigenschaften, die wir von Kontraktionen nicht generell erwarten. So beseitigt eine Kontraktion nach vollem Schnitt eine Konjunktion $A \wedge B$, indem sie immer *beide* Konjunkte beseitigt, d.h. solche Kontraktionen haben generell die Eigenschaft (*intersection* in [5, §6])

$$\text{(CI)} \qquad \forall A, B \in K: \; K - A \wedge B = K - A \cap K - B.$$

Tatsächlich ist diese Eigenschaft charakteristisch für den vollen Schnitt, denn es gilt auch umgekehrt, daß eine Operation mit den Eigenschaften (C1-6) und (CI) als Kontraktion nach vollem Schnitt darstellbar ist (Beobachtung 6.4 in [5]). Des weiteren haben Vollschnittkontraktionen auch diese Eigenschaft (siehe [7]):

$$\text{(C}\dagger\text{)} \qquad \text{Wenn } A \in K, \text{ dann } K - A = K \cap \mathrm{Cn}(\neg A).$$

Damit werden Kontraktionen sicher zu klein. Wenn nun Revisionen über die (LI) definiert werden, dann enthält $K * A$ (im Hauptfall $\neg A \in K$) nur noch die Konsequenzen von A, d.h.

$$\text{(R}\dagger\text{)} \qquad \text{Wenn } \neg A \in K, \text{ dann } K * A = \mathrm{Cn}(\neg A).$$

Während der volle Schnitt zu drastisch in ein Überzeugungssystem eingreift, führt die Maximalwahl zu einer anderen Art ungewollter, nämlich zu starker Eigenschaften. So erfüllen Kontraktionen nach Maximalwahl die folgende Bedingung (siehe [7]):

$$\text{(C}\ddagger\text{)} \qquad \text{Wenn } A \in K, \text{ dann } A \vee B \in K - A \text{ oder } A \vee \neg B \in K - A.$$

Wenn wir nun aus einer Kontraktion nach Maximalwahl eine Revision mittels der (LI) generieren, dann ist (wieder im Hauptfall) das Resultat, $K * A$, vollständig, d.h.

$$\text{(R}\ddagger\text{)} \qquad \text{Wenn } \neg A \in K, \text{ dann } B \in K * A \text{ oder } \neg B \in K * A.$$

Im weiteren werden uns voller Schnitt und Maximalwahl nicht weiter beschäftigen. Kontraktion nach Teilschnitt scheint genau die Goldene Mitte zwischen diesen Extremen zu treffen.

Bevor wir uns den Repräsentationsresultaten zuwenden, sei darauf hingewiesen, daß, gleichgültig wie zwei Mengen zueinander stehen (z.B. $K \subseteq H$), die AGM-Theorie nichts über das Verhältnis zwischen $s(K \perp A)$ und $s(H \perp A)$ sagt. Im ersten Ausdruck ist s eine Auswahlfunktion, die für K zuständig ist, im zweiten Ausdruck ist s eine *andere* Auswahlfunktion, diesmal für H zuständig. (Um diese Abhängigkeit der Auswahl vom betrachteten Überzeugungszustand deutlich zu machen, schreiben wir manchmal s_K bzw. s_H.) Es sei z.B. X eine Menge, die von s_H in $H \perp A$ ausgewählt wird und zugleich in $K \perp A$ ist. Dann nehmen wir hier *nicht* an, daß auch s_K die Menge X in $K \perp A$ auswählen muß. Ebenso: Wenn wir uns die Auswahl s_K durch eine Präferenzordnung $\leq_K$ bestimmt vorstellen, dann nehmen wir *nicht* an, daß mit einer Erweiterung von K zu H nur eine bloße Erweiterung der Ordnung $\leq_K$ einhergehen kann.

Einerseits ist diese Unabhängigkeit der Auswahlfunktion von den jeweiligen Überzeugungen aus zwei Gründen sicherlich ein Vorzug der Theorie. Erstens, können zwei Subjekte mit identischen Überzeugungen auf Evidenz verschieden reagieren, ohne daß eines der beiden deshalb unvernünftig sein muß. Die zwei Subjekte können einfach verschiedene epistemische Nutzenfunktionen haben – unterschiedliches "intellektuelles Temperament", wie William James sagen würde. Zweitens, muß die Lösung einer Auswahlaufgabe auch abhängig sein von der Überzeugungsmenge, in der die Auswahl getroffen werden soll. So betrachtet, gibt es keinen Grund zu fordern, daß die Auswahl, die für eine Menge K getroffen wird, die Auswahl für eine Obermenge H immer irgendwie einschränken muß. Denkökonomie mag so etwas empfehlen. Aber Denkökonomie ist ein übertrumpfbares Vernunftgebot.

Andererseits führt die Unabhängigkeit der zwei Komponenten eines Überzeugungssystems – der Überzeugungsmenge und der Auswahlfunktion – zu einem Problem der "kategorialen Passung"; siehe [103, p. 37] und [251, §3.6]. Gegeben eine Überzeugungsmenge und eine Auswahlfunktion, bestimmt eine Teilschnittkonstruktion für jede Aussage eine neue Überzeugungsmenge – jedoch keine neue Auswahlfunktion. Die Teilschnittkonstruktion für eine zu entfernende Aussage bildet also ein Überzeugungs*system* in eine Überzeugungs*menge* ab. Das Argument der Funktion paßt nicht zu ihrem Wert. Es fehlt hier eine Komponente, die insbesondere nötig wäre, um weitere Überzeugungsänderungen durchzuführen. So führt das Problem der kategorialen Passung direkt zum Problem iterierter Überzeugungsänderungen.

Es ist allerdings nicht so, daß iterierte Überzeugungsänderungen nicht Gegenstand der AGM-Theorie sein könnten. Iterierte Kontraktionen wie $(K - A) - B$ fallen ohne weiteres in den Skopus der AGM-Postulate und sind auch über Teilschnitte oder äquivalente Konstruktion wohldefiniert. Es ist nur so, daß keine logische Beziehung zwischen s_K und s_{K-A} angenommen wird. Bei jeder Auswahlaufgabe können die Karten gewissermaßen neu gemischt werden. Die kategoriale Passung zwischen Ausgangspunkt K und Resultat $K - A$ einer Kontraktion wird wieder hergestellt, indem die neue Überzeugungsmenge mit einer *beliebigen* neuen Auswahlfunktion s_{K-A} zu einem neuen Überzeugungssystem ergänzt werden kann.

Wir schließen diesen Abschnitt mit einem Lemma über Restefamilien, welches wir für die Repräsentationsresultate benötigen werden.

LEMMA 8. *Wenn $A \in K$, $X \in K\bot B$, und $X \nvdash A$, dann $X \in K\bot A$.*

BEWEIS. Wir nehmen an, daß (1) $A \in K$, (2) $X \in K\bot B$, und (3) $X \nvdash A$. Aus (2) folgt, daß $X \subseteq K$. Gegeben (3), genügt es für $X \in K\bot A$ zu zeigen, daß $\forall Y : X \subset Y \subseteq K \Rightarrow Y \vdash A$. Wir nehmen also an, daß

$$X \subset Y \subseteq K \tag{4}$$

und zeigen, daß $Y \vdash A$. — Aus (2) folgt, daß $K\bot B \neq \emptyset$ und daher ist (ebenfalls nach 2)

$$\bigcap(K\bot B) \subseteq X. \tag{5}$$

Nun gilt nach Teilsatz 1 Zurück für jede Wahl von σ in $\bigcap\sigma(K\bot B) \subseteq X$ (siehe den Beweis oben), also auch für den Fall $\sigma(K\bot B) = K\bot B$. Daher:

$$K \subseteq \mathrm{Cn}(\bigcap(K\bot B) \cup \{B\}). \tag{6}$$

Es folgt aus (6), (5) und (4), daß

$$K \subseteq \mathrm{Cn}(Y \cup \{B\}). \tag{7}$$

Nun folgt aber auch aus (2) und (4) $X \subset Y$, daß $Y \vdash B$. Also ist $\mathrm{Cn}(Y \cup \{B\}) = \mathrm{Cn}(Y)$. So folgt aus (1) und (7) wie gewünscht, daß $Y \vdash A$. ∎

Repräsentationsresultat für die Basispostulate. Wir zeigen zunächst, daß die Kontraktionspostulate C1–6 genau die Teilschnittkontraktionen beschreiben. Sodann definieren wir Teilschnittrevisionen nach der (LI) und schließen dann aufgrund von Beobachtung 9, daß diese genau durch die Revisionspostulate R1–6 beschrieben werden. Danach wenden wir uns den Zusatzpostulaten 7 und 8 zu.

Satz 9. Repräsentation für C1-6 (AGM [5, §2]).

1. *Jede Teilschnittkontraktion erfüllt die Bedingungen C1-6.*
2. *Jede Operation $-$, welche C1-6 erfüllt, kann durch eine Teilschnittkontraktion dargestellt werden.*

Beweis. 1. *Jede Teilschnittkontraktion erfüllt C1-6.*
Wir illustrieren den Beweis mit Zurück: $K \subseteq K - A + A$. Es ist also zu zeigen, daß $K \subseteq \mathrm{Cn}(\bigcap s(K\perp A) \cup A)$, d.h. unter der Annahme $K \vdash B$ zeigen wir, daß $\bigcap s(K\perp A), A \vdash B$.

Zunächst betrachten wir einen Grenzfall: Wenn $K \nvdash A$, dann ist $K\perp A = \emptyset$ und so ist $s(K\perp A) = \{K\}$ – fertig. Weiter mit dem Hauptfall $K \vdash A$. Dann gilt:

$$K \cap \mathrm{Cn}(\neg A) \subseteq \bigcap(K\perp A). \tag{$*$}$$

(Denn, angenommen (1) $B \in K$ und (2) $\neg A \vdash B$, und – für *reductio* – $B \notin \bigcap(K\perp A)$. Da $A \in K$, so gibt es ein $X \in K\perp A$ mit $B \notin X$ und also (nach 1) $X, B \vdash A$. Wir haben aber auch $\neg B \vdash A$ aus (2). Also $X \vdash A$ – im Widerspruch zu unserer Annahme, daß $X \in K\perp A$. Die Inklusion $(*)$ gilt übrigens auch in umgekehrter Richtung – vgl. (C†) oben –, was uns an dieser Stelle aber nicht interessiert.)

Die gewünschte Konklusion erreichen wir nun so:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{K\vdash B}{K\vdash A\to B}\qquad \dfrac{}{\neg A\vdash A\to B}}{K\cap \mathrm{Cn}(\neg A)\vdash A\to B}\qquad (*)}{\bigcap(K\perp A)\vdash A\to B}\qquad \dfrac{}{\bigcap(K\perp B)\subseteq\bigcap s(K\perp B)}}{\bigcap s(K\perp A)\vdash A\to B}}{\bigcap s(K\perp A), A\vdash B}$$

2. *Jede (C1-6)-Kontraktion läßt sich als Teilschnittkontraktion darstellen.*
Wir nehmen an, $K - (\ )$ sei eine Kontraktion (für eine beliebig gewählte Theorie K), welche C1-6 erfüllt. Wir definieren eine kanonische Auswahlfunktion $\sigma_{(K)}$ für alle Formeln A:

$$\sigma(K\perp A) = \begin{cases} \{X \in K\perp A : K - A \subseteq X\}, & \text{falls } K\perp A \neq \emptyset; \\ \{K\}, & \text{anderenfalls.} \end{cases} \tag{σ}$$

Nun ist zu zeigen, daß σ die Bedingungen für s in Definition 7 erfüllt; im einzelnen:

(1) σ ist wohldefiniert, d.h. aus $K\bot A = K\bot B$ folgt $\sigma(K\bot A) = \sigma(K\bot B)$.
(2) $\sigma(K\bot A) = \{K\}$, falls $K\bot A = \emptyset$.
(3) $\sigma(K\bot A) \subseteq K\bot A$, falls $K\bot A \neq \emptyset$.
(4) $\sigma(K\bot A) \neq \emptyset$.
Und schließlich:
(5) $K - A = \bigcap \sigma(K\bot A)$.

Ad (1–4). (1) folgt aus Kongruenz (C6). (2) und (3) folgen unmittelbar aus (σ). Im Grenzfall $K\bot A = \emptyset$, d.h. $\vdash A$, ist (4) nach (σ) erfüllt. Anderenfalls haben wir $\nvdash A$ und so gilt nach Erfolg (C2) $K - A \nvdash A$. Dann $\exists X \supseteq K - A$ mit $X \in K\bot A$ und wieder ist (4) erfüllt.

Ad (5). Die Beziehung $K - A \subseteq \bigcap \sigma(K\bot A)$ ergibt sich unmittelbar aus (σ) (im Falle $K\bot A = \emptyset$) aus Schranke (C3). Es bleibt zu zeigen:

$$(\dagger) \qquad \text{Wenn } B \notin K - A, \text{ dann } B \notin \bigcap \sigma(K\bot A).$$

Fall $A \notin K$. Dann ist $K\bot A = \emptyset$ und also ist $\bigcap \sigma(K\bot A) = K$. Ferner ist nach Null (C4) $K \subseteq K - A$. So folgt unter der Annahme $B \notin K - A$, daß $B \notin K$ $[= \bigcap \sigma(K\bot A)]$.

Fall $A \in K$. Nun ist $\sigma(K\bot A) = \{X \in K\bot A : K - A \subseteq X\}$. ($\dagger$) ist trivialerweise wahr, wenn $B \notin K$. Also nehmen wir ferner an, $B \in K$. Es gilt jetzt eine Menge X zu finden so, daß

$$\text{(a) } K - A \subseteq X \quad \text{und} \quad \text{(b) } X \in K\bot A \quad \text{und} \quad \text{(c) } B \notin X.$$

Ad a)
$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{B \in K}\,\text{Ann.}}{K - A, A \vdash B}\,\text{Zurück} \qquad \dfrac{}{K - A \nvdash B}\,\text{Ann.}}{K - A, \neg A \nvdash B}}{K - A \nvdash A \vee B}}{\exists X \in K\bot(A \vee B) : K - A \subseteq X}\,\text{Def.}\bot$$

Für (b) verwenden wir Lemma 8, welches hier in dieser Form in Erinnerung gerufen sei:

$$\text{Lemma} \qquad \frac{C \in K \quad X \in K\bot D \quad X \nvdash C}{X \in K\bot C}.$$

Ad b)

$$\dfrac{\dfrac{}{A \in K}\,\text{Ann.} \qquad \dfrac{}{X \in K\bot(A \vee B)}\,\text{(a)} \qquad \dfrac{\dfrac{\dfrac{}{X \in K\bot(A \vee B)}\,\text{(a)}}{X \nvdash A \vee B}}{X \nvdash A}}{X \in K\bot A}\,\text{Lemma}$$

Ad c)

$$\dfrac{\dfrac{\dfrac{}{X \in K\bot A \vee B}\,\text{(a)}}{X \nvdash A \vee B}}{B \notin X}$$

■

Teilschnittrevisionen wollen wir so definieren:

$$K * A = \mathrm{Cn}(\bigcap s(K\bot\neg A) \cup \{A\}).$$

Diese Definition ist nichts anderes als die (LI) mit der Teilschnittkontraktion $K - \neg A = \bigcap s(K\bot\neg A)$ im *definiens*. Da Teilschnittkontraktionen die Postulate C1–6 erfüllen (Satz 9.1), dürfen wir aus Beobachtung 2 schließen, daß Teilschnittrevisionen die Postulate R1–6 erfüllen. Umgekehrt, wenn eine Operation $*$ die Bedingungen R1–6 erfüllt, dann ist sie nach Beobachtung 2 mittels der (LI) durch eine Kontraktion darstellbar. Diese sind aber wiederum (nach 9.2) als Teilschnittkontraktionen darstellbar. D.h., wenn $*$ die R-Bedingungen erfüllt, dann ist $*$ als Teilschnittrevision im Sinne der gerade eingeführten Definition darstellbar. So haben wir den folgenden Satz bewiesen:

Satz 10. Repräsentation für R1-6 (AGM [5]). *Eine Operation $*$ erfüllt genau dann die Bedingungen (R1–6), wenn sie als eine Teilschnittrevision darstellbar ist.*

Repräsentationsresultat für die Zusatzpostulate. Wir kommen nun zu den Zusatzpostulaten 7 und 8. Zwar haben wir oben (p. 509) davon gesprochen, daß eine Wahl zwischen Kontraktionskandidaten eine Präferenz voraussetzt, jedoch ist Präferenz bisher nicht im Sinne einer Ordnung von Optionen zur Darstellung gekommen. Vielmehr haben wir Präferenzen indirekt mithilfe einer Auswahlfunktion dargestellt.

Es sei P eine Menge von Optionen. Dann können wir sagen, daß eine Auswahl s in P auf einer Relation $\leq\, \subseteq P \times P$ basiert, wenn $s(P)$ aus den "besten" Elementen (den größten unter $\leq$) in P besteht:

$$s(P) = \{x \in P : (\forall y) \text{ wenn } y \in P, \text{ dann } y \leq x\}.$$

Allein schon die Annahme, daß die Auswahl in einer Restefamilie auf einer Relation basiert, genügt, damit eine Teilschnittkontraktion (C7) erfüllt. Es sei

$$\mathcal{T}(K) = \{K' \subseteq K : \mathrm{Cn}(K') \subseteq K'\}$$

die Menge aller Teiltheorien von K. Wir wollen eine Teilschnittkontraktion $-$ *relational* nennen, wenn es für jede Theorie K eine Relation $\leq\, \subseteq \mathcal{T}(K) \times \mathcal{T}(K)$ gibt so, daß für alle Formeln A die Auswahl s in $K - A = \bigcap s(K\bot A)$ auf $\leq$ basiert, d.h.

$$(s_{\leq}) \qquad K - A = \bigcap\{X \in K\bot A : (\forall Y)\ Y \in K\bot A \Rightarrow Y \leq X\}.$$

Für den folgenden Satz beweisen wir zunächst ein weiteres Lemma über Restefamilien.

Lemma 11. *Wenn $A, B \in K$, dann $K\bot(A \wedge B) = K\bot A \cup K\bot B$.*

Beweis. $\subseteq$: Wir nehmen an, daß (1) $A, B \in K$ und (2) $X \in K\bot(A \wedge B)$. Es folgt, daß (3a) $X \nvdash A$ oder (3b) $X \nvdash B$. Falls (3a), dann schließen wir nach Lemma 8 aus (1) und (2), daß $X \in K\bot A$; falls (3b), dann ebenso $X \in K\bot B$. In beiden Fällen also $X \in K\bot A \cup K\bot B$.

$\supseteq$: Angenommen (1) $A, B \in K$ (also auch $A \wedge B \in K$) und (2a) $X \in K\bot A$ oder (2b) $X \in K\bot B$. In beiden Fällen haben wir (3) $X \nvdash A \wedge B$. In beiden Fällen können wir aus (1) und (3) nach Lemma 8 auf daß $X \in K\bot(A \wedge B)$ schließen. ■

Satz 12. *Jede relationale Teilschnittkontraktion erfüllt die Bedingung (C7).*

Beweis. Falls $A \notin K$ oder $B \notin K$ ist der Satz trivial wahr. Wir betrachten also den Fall

$$(1) \qquad A \in K \text{ und } B \in K.$$

Nach einfacher Mengenumformung ist zu zeigen, daß

$$\bigcap(s(K\bot A) \cup s(K\bot B)) \subseteq \bigcap s(K\bot(A \wedge B)).$$

Dafür genügt der Nachweis, daß

$$s(K\bot(A \wedge B)) \subseteq s(K\bot A) \cup s(K\bot B),$$

was nach (1) und Lemma 11 äquivalent ist zu

$$s(K\bot A \cup K\bot B) \subseteq s(K\bot A) \cup s(K\bot B).$$

Wir nehmen an, daß

(2) $$X \in s(K\bot A \cup K\bot B)\ [\subseteq K\bot A \cup K\bot B].$$

Also (a) $X \in K\bot A$ oder (b) $X \in K\bot B$. Die beiden Fälle sind völlig gleich.

Wir behandeln Fall (a). Wir betrachten ein beliebiges $Y \in K\bot A$ [$\subseteq K\bot A \cup K\bot B$]. Es folgt aus Annahme (1) und Lemma 11, daß $Y \in K\bot(A \wedge B)$. Nach (1–2) und dem Lemma ist $X \in s(K\bot(A \wedge B))$. Da s relational über $\leq$ in $\mathcal{T}(K)$ ist, so muß $Y \leq X$ sein. So haben wir gezeigt, daß $(\forall Y)\ Y \in K\bot A \Rightarrow Y \leq X$, d.h. $X \in s(K\bot A)\,[\cup\, s(K\bot B)]$. ∎

Über die Relationen $\leq\ \subseteq \mathcal{T}(K)^2$ haben wir bisher nichts weiter angenommen. Aber der Idee nach soll es sich dabei um eine Präferenzordnung handeln. So kommen gleich eine Reihe formaler Eigenschaften ins Spiel, zu denen insbesondere

Transitivität: wenn $x \leq y$ und $y \leq z$, dann $x \leq z$.

gehört. Eine Teilschnittkontraktion – heiße *transitiv-relational*, wenn es für jede Theorie K eine transitive Relation in $\mathcal{T}(K)$ gibt so, daß ($s_{\leq}$), für jede Formel A.

Satz 13. *Jede transitiv-relationale Teilschnittkontraktion erfüllt die Bedingung (C8).*

Beweis. Wir nehmen an, daß

(1) $$A \notin \bigcap s(K\bot(A \wedge B))$$

und, wie im vorigen Beweis, daß (2) $A \in K$ und $B \in K$ (anderenfalls gilt der Satz wieder trivial). Es wird genügen, zu zeigen, daß $s(K\bot A) \subseteq s(K\bot(A \wedge B))$. Aus (1) folgt, daß

(3) $$K\bot A \cap s(K\bot(A \wedge B)) \neq \emptyset.$$

(Denn nach (1) gibt es ein $X \in s(K\bot(A \wedge B)) \subseteq K\bot(A \wedge B)$ mit $A \notin X$. Da aber (2!) $A \in K$, so können wir Lemma 8 anwenden: $X \in K\bot A$.)

Nun nehmen wir – für *reductio* – an, daß es ein X gibt so, daß (4) $X \in s(K\bot A)$ und (5) $X \notin s(K\bot(A \wedge B))$. Aus (2) folgt nach Lemma 11,

$$s(K\bot A)\ \subseteq\ K\bot A \cup K\bot B\ \subseteq\ K\bot(A \wedge B).$$

Also folgt aus (4), daß (6) $X \in K\bot(A \wedge B)$, jedoch (5!) $X \notin s(K\bot(A \wedge B))$. Daher gibt es ein (7) $Y \in K\bot(A \wedge B)$ mit (8) $Y \not\leq X$. Nach (3) gibt es ein (9) Z in $K\bot A$ mit (10) $Z \in s(K\bot(A \wedge B))$. Nun folgt aus (4) und (9), daß $Z \leq X$, und aus (7) und (10), daß $Y \leq Z$. Aber dann per Transitivität $Y \leq X$ – im Widerspruch zu (8). ∎

Schließlich können wir die letzten zwei Sätze zu einem Repräsentationsresultat für (C1–8)-Kontraktionen erweiterm

Satz 14. Repräsentation für (C1-8) (AGM [5, §4]). *Eine Operation – erfüllt genau dann die Bedingungen (C1–8), wenn sie eine transitiv-relationale Teilschnittkontraktion ist.*

Beweis. (Skizze.) Die eine Hälfte haben wir oben bewiesen. Für die andere Hälfte verfahren wir so wie im Beweis von Satz 9. $K - (\ )$ sei eine Kontraktion für eine beliebig gewählte Theorie K so, daß die Bedingungen (C1–8) erfüllt sind. Wir definieren eine Relation $R \subseteq \mathcal{T}(K)^2$:

$$(R)\quad XRY \text{ gdw } \begin{cases} X = K = Y;\ \text{oder:} \\ \text{(a) } \exists A \in K : X \in K\bot A, \text{ und} \\ \text{(b) } \exists B \in K : Y \in K\bot B \text{ und } K - B \subseteq Y, \text{ und} \\ \text{(c) } \forall C : X, Y \in K\bot C \ \&\ K - C \subseteq X, \Rightarrow K - C \subseteq Y. \end{cases}$$

Nun ist zu zeigen, daß R eine transitive Relation ist und für $\leq$ in die Bedingung $(s_{\leq})$ eintreten kann. Im einzelnen also:

(1) R ist wohldefiniert, d.h. aus $\mathrm{Cn}(X) = \mathrm{Cn}(X')$ folgt, XRY falls $X'RY$, und YRX falls YRX'.

(2) R ist transitiv.

(3) $K - A = \bigcap\{X \in K\bot A : (\forall Y)\ Y \in K\bot A \Rightarrow YRX\}$.

Der Beweis ist in [5, §4] und in [100, App. B] ausgeführt. ■

Genauso wie wir in Satz 10 den einfachen Repräsentationssatz 9 für Kontraktionen unter den Grundpostulaten (C1–6) auf Revisionen übertragen haben, so können wir nun den erweiterten Repräsentationssatz 14 für (C1–8)-Kontraktionen auf Revisionen übertragen. Danach erfüllt eine Operation $*$ genau dann die Bedingungen (R1–8), wenn sie eine transitiv-relationale Teilschnittrevision ist.

7. Drei weitere Modellierungen

Mit der nötigen Vorsicht, die Beobachtung 4 anmahnt, gelingt es über den Ramsey-Test die üblichen Konditionallogiken hervorzubringen (siehe Satz 6). Die vorsichtige Variante der AGM-Theorie stellt also so etwas wie eine Semantik kontrafaktischer Konditionale dar. Das wirft die Frage auf, in welcher Beziehung die AGM-Theorie zu anderen Arten solche Konditionale zu modellieren, steht. Die Sphärensysteme von Lewis [182] sind eine besonders prominente Modellierung, deren Beziehung zu alternativen Darstellungen überdies eingehend erforscht ist. Wir werden in diesem Abschnitt darstellen, wie Sphärenmodelle adaptiert werden können, um Revisionen und Kontraktionen zu konstruieren. Zuvor stellen wir kurz zwei weitere Konstruktionen vor.

Sichere Kontraktionen. Wollen wir eine Überzeugung A aufgeben, so fragen wir oft danach, wie sie in anderen unserer Überzeugungen begründet ist. Solche Überzeugungen, die in keinem Begründungsverhältnis zu A stehen, müssen wir nicht aufgeben, wenn wir A aufgeben wollen: Sie sind sicher im Hinblick auf A. Aber auch wenn ein Teil X unserer Überzeugungen A stützt, so müssen wir nicht alles in X aufgeben. Es genügt, in X sichere von unsicheren Annahmen zu unterscheiden und nur letztere zurückzuziehen. Am Ende sollten wir bei den Überzeugungen bleiben, die vergleichsweise sicher sind. Das ist die Grundidee der sogenannten *sicheren Kontraktion.*[15] Sie wurde erstmals von Alchourrón und Makinson in [6] vorgestellt und in [7, 8] weiter untersucht. Die Resultate von Alchourrón und Makinson konnten Rott und Hansson in [252] bedeutend erweitern.

Für "Begründungsverhältnis" lesen wir logisches Folgerungsverhältnis. Angenommen $A \in K$ und wir wollen A aus K entfernen. Dann bilden wir zunächst die *minimalen Prämissenmengen in K für A*, d.h. alle solche Teilmengen von K, aus denen A folgt, ohne daß A aus einem Teil davon folgt.

$$X \in K\angle A \text{ gdw } X \subseteq K, \text{ und } X \vdash A, \text{ und } \forall Y : Y \subset X \Rightarrow Y \nvdash A.$$

Die Kompaktheit der Relation $\vdash$ garantiert, daß alle Mengen in $K\angle A$ endlich sind. Sodann führen wir eine *Relation $< \subseteq K \times K$ relativer Sicherheit* ein. Für $A, B \in K$ soll $A < B$ ausdrücken, daß B sicherer als A in K ist. Wir fordern hier nur, daß $<$ in jeder endlichen Teilmenge von K – also insbesondere in $K\angle A$-Mengen – minimale (unsicherste) Elemente auszeichnet. (B ist *minimal* unter $<$ in X, wenn es kein $C \in X$ gibt mit $C < B$.) Dafür genügt die Bedingung, daß $<$ in K azyklisch sei, d.h. daß es in K keine Kette $A_0 < \cdots < A_n$ gebe mit $A_n = A_0$. Wenn ein Element B in einer Menge $X \in K\angle A$ minimal ist unter $<$, dann ist B gewissermaßen "schuld" an der Ableitbarkeit von A und muß in der Folge seinen Hut nehmen.[16]

Jetzt können wir sagen, daß eine Überzeugung B in K *sicher im Hinblick auf A* ist, wenn B entweder in keiner Menge in $K\angle A$ vorkommt oder B in keiner Menge $X \in K\angle A$, in der es vorkommt, $<$-minimal ("schuldig") ist.

[15] Es ist auch die Grundidee von Hanssons *kernel*-Kontraktion in [127], vgl. [129, §2.8]. Siehe auch die nächste Fußnote.

[16] Wenn $<$ diese Bedingung erfüllt, dann nennen Alchourrón und Makinson [6] $(K, <)$ eine "*Hierarchie* über K".

Hanssons [127] *kernel*-Kontraktion verfährt etwas allgemeiner, indem nur angenommen wird, daß es für jedes K eine Auswahlfunktion über $\bigcup K \perp A$ $(\forall A)$ gibt, welche die zu entfernenden Schuldigen benennt. Sichere Kontraktionen sind dann *kernel*-Kontraktionen, in denen die Auswahl durch eine azyklische Relation bestimmt ist. *Kernel*-Kontraktionen stehen also zu sicheren Kontraktionen ähnlich wie allgemeine zu relationalen Teilschnittkontraktionen.

Die Kontraktion $K - A$ soll dann aus allen Überzeugungen in K bestehen, die sicher im Hinblick auf A sind (genauer: aus dem, was daraus folgt). Wir fassen das in der folgenden Definition zusammen.

Definition 15. Es sei K eine Theorie und $< \subseteq K^2$ sei azyklisch. Für jede Formel A sei $K \angle A$ wie oben definiert.

1. *Sichere Elemente* in K im Hinblick auf A:

$$Si(K\angle A) = \{B : B \in K \text{ und } \forall X \in K\angle A : B \in X \Rightarrow \exists C \in X : C < B\}.$$

2. $K - A$ ist genau dann eine *sichere Kontraktion* über K, wenn

$$K - A = \mathrm{Cn}(Si(K\angle A)).$$

Was geschieht, wenn $<$ leer ist? In diesem Fall sind alle Elemente einer Menge in $K\angle A$ minimal unter $<$. Sicher sind dann nur solche Elemente in K, die in keiner Menge in $K\angle A$ vorkommen.

Sichere Kontraktionen sind ganz anders konstruiert als Teilschnittkontraktionen. Vor eine Kontraktionsaufgabe $K - A$ gestellt, betrachten wir im einen Fall $\subseteq$-minimale Mengen, die A implizieren, im anderen Fall $\subseteq$-maximale Menge, die A ausschließen. Im einen Fall machen wir eine Bestenauslese unter Teilmengen von K, im anderen Fall eine Schlechtestenauslese in K selbst. Dennoch lassen sich sichere Kontraktionen unter die Klassen der Teilschnittkontraktionen subsumieren. Dies folgt nach Satz 9.2 unmittelbar aus dem folgenden Satz.

Satz 16. (Alchourrón und Makinson [9].) *Jede sichere Kontraktion erfüllt die AGM-Postulate (C1–6).*

Beweis. Es sei S kurz für $Si(K\angle A)$.

(C1) und (C3) sind unmittelbar gegeben. Für (C4) beobachten wir, daß $K\angle A = \emptyset$, wenn $A \notin K$. Daher $K = S$. (C6) folgt aus dem Umstand, daß $K\angle A = K\angle B$, falls $A \equiv B$.

Ad Erfolg (C2): Für *reductio* nehmen wir an, daß $S \vdash A$. Dann gibt es eine nichtleere Menge $X \subseteq S$ in $K\angle A$. Da $\nvdash A$, so ist $X \neq \emptyset$. Sei B also ein minimales Element in X. Dann $B \notin S$. Aber $B \in X \subseteq S$ – Widerspruch.

Ad Zurück (C5): Wir beweisen zunächst, daß

$$(*) \qquad K \cap \mathrm{Cn}(\neg A) \subseteq S.$$

Angenommen (1) $K \vdash B$, (2) $\neg A \vdash B$ und (3) $S \nvdash B$. Aus (1) und (3) folgt, daß (4) B minimal ist in einer Menge $X \in K\angle A$. Wir argumentieren nun so:

$$\dfrac{\dfrac{(4)}{X \setminus \{B\}, B \vdash A} \qquad \dfrac{(2)}{\neg B \vdash A}}{\dfrac{X \setminus \{B\}, B, \neg B \vdash A}{X \setminus \{B\} \vdash A.}}$$

Aber dann kann X (mit B) nicht in $K \angle A$ sein – Widerspruch!

Angenommen nun, $K \vdash B$. Dann $K \vdash A \to B$. Aber wir haben auch $\neg A \vdash A \to B$. So folgt nach $(*)$ $B \in S \subseteq \mathrm{Cn}(S)$, wie gewünscht. ■

Die Umkehrung des Satzes gilt nicht! Rott und Hansson [252, Theorem 6] führen ein Gegenbeispiel vor. Das Beispiel ist eine Kontraktionsfunktion über einer Theorie K, welche einerseits für bestimmte Input-Formeln Werte ausgibt, die mit den Bedingungen (C1–C6) vereinbar sind. Diese Werte müssen also nach Satz 9 über Teilschnittkonstruktionen erzeugt werden können. Andererseits können diese Werte x nur dann in die Gleichung $x = \mathrm{Cn}(Si(K \angle A))$ eingesetzt werden, wenn die Relation $<$, die Si zugrundeliegt, zyklisch ist. Also gibt es Kontraktionen, die durch Teilschnitte, aber nicht durch (azyklische) Sicherheitsrelationen erzeugt werden können.

Besser stehen die Dinge, wenn wir ein Repräsentationsresultat für die Postulate (C1–8) anstreben und dazu die Bedingung der Azyklizität von $<$ um weitere ergänzen. Drei Bedingungen spielen hier eine Rolle. Die erste fordert, daß sich $<$ in einer Theorie K aufwärts über $\vdash$ fortsetzt:

(<Auf) Wenn $A < B$ und $B \vdash C$, dann $A < C$.

Unter naheliegenden Interpretationen der Relation $<$ als relativer Sicherheit klingt das sehr plausibel. Insbesondere, wenn wir annehmen, daß logisch stärkere Aussagen immer unsicherer sind als die schwächeren, die aus ihnen folgen, dann dürfen wir aus $B \vdash C$ auf $B < C$ schließen und (Auf) entpuppt sich als eine Art Transitivität. Aber in dieser Überlegung nehmen wir schon Eigenschaften von $<$ an, die zusammen stärker sind als (Auf). Wir können auch etwas anders für (Auf) argumentieren. Dazu erinnern wir uns, daß $<$ allein die Aufgabe hat, "Schuldige" zu identifizieren. Wenn $A < B$, dann ist A ein potentieller Schuldiger. Wenn ferner auch C minimal ist, dann könnten beide, A und C, schuldig gesprochen werden. Aber wenn nun C aus B folgt, dann muß mit C auch B aufgegeben werden. Also könnten wir ohne (Auf) gezwungen sein, alle drei Aussagen zurückzuziehen. (Auf) schließt solche Situationen aus, indem es unter den beiden Annahmen C aus dem Kreis der Verdächtigen entläßt. So gesehen ist (Auf) ein Mittel, die Maxime des minimalen Eingriffs umzusetzen. Ähnlich können wir für die folgende Bedingung argumentieren.

(<Ab) Wenn $A \vdash B$ und $B < C$, dann $A < C$.

Die dritte zu betrachtende Bedingung, virtuelle Konnexität, ist in unserem Zusammenhang rein technischer Natur:

(<VKon) Wenn $A < B$, dann $A < C$ oder $C < B$.

Die stärkere Bedingung $A < B$ ist nötig, da die starke Form ($A \neq B \Rightarrow A < B$ oder $B < A$) keine azyklische Relation mit (Auf) und (Ab) erfüllen kann. (Man betrachte zwei Formeln $A \neq B$ so, daß $A \vdash B$ und $B \vdash A$. Wir hätten nun $A < B$ oder $B < A$. Aber dann erhalten wir im einen Fall $A < A$, im anderen $B < B$ nach (Auf) bzw. (Ab); d.h. $<$ wäre zyklisch.)

Wenn $<$ azyklisch ist, dann ist $<$ offensichtlich auch irreflexiv. Es ist ferner leicht nachzuprüfen, daß wenn $<$ irreflexiv und virtuell konnex ist, dann ist $<$ transitiv und (Auf) und (Ab) sind äquivalent.

Zum Beispiel (VKon) & (Irr) & (Auf) $\Rightarrow$ (Ab): Angenommen (1) $B < C$. Dann nach (VKon) (a) $C < A$ oder (b) $A < C$. Im Fall (b) folgt die gewünschte Konklusion ohne die weitere Annahme (2) $A \vdash B$. Fall (b) können wir ausschließen. Denn aus (1) und (a) folgt per Transitivität ($\Leftarrow$ VKon & Irr), daß $B < A$. Aber dann aus (2) nach (Auf), $A < A$ – im Widerspruch zu (Irr).

Nun können wir folgendes beobachten (siehe [9]):

- Jede sichere Kontraktion unter der Bedingung (Auf) oder (Ab) erfüllt (C7).
- Jede sichere Kontraktion unter der Bedingung (Auf) oder (Ab), sowie (VKon) erfüllt (C8).

Also erfüllt jede sichere Kontraktion einer Theorie K, in der die Relation $<_K$ virtuell konnex ist und sich aufwärts oder abwärts über logische Folgerung fortsetzt, alle AGM-Postulate. Da auch die Umkehrung gilt (was wir hier aber nicht beweisen werden), dürfen wir festhalten:

Satz 17. (Rott [250].) *Eine sichere Kontraktion (über K) erfüllt genau dann die AGM-Postulate (C1–8), wenn $<_{(K)}$ die Bedingungen (Auf) oder (Ab), sowie (VKon) erfüllt.*

Zusammen mit Satz 9.2 folgt so, daß sichere Kontraktionen unter den im Satz genannten Bedingungen mit dem Ergebnis relationaler Teilschnitte übereinstimmen. Die beiden Konstruktionen sind also gleichwertig.

Epistemische Verankerung. Die nun folgende Konstruktion ist mit der gerade behandelten in zweierlei Hinsicht eng verwandt. Sie taucht erstmals in Gärdenfors' [98] unter der Bezeichnung *epistemic entrenchment* auf und wird eingehender in [101] und [100] behandelt. In Rott [251] (und weiteren Publikationen) wird die Theorie erweitert und in ihren Querverbindungen zu anderen Konstruktionen untersucht.

Die Verwandtschaft mit sicheren Kontraktionen besteht, erstens, darin, daß die jetzige Konstruktion ebenfalls mit einer irreflexiven Relation $\prec \subseteq K \times K$ beginnt. Wir werden im folgenden sagen, die Relation vergleiche die Elemente eines Überzeugungssystems K nach ihrer *epistemischen Verankerung.* Die Verwandtschaft zwischen epistemischer Verankerung und Sicherheit setzt sich, zweitens, darin fort, daß die formalen Eigenschaften dieser beiden Relationen weitgehend übereinstimmen, wie wir sehen werden. [17] Die jeweiligen Konstruktionen von Kontraktionen unterscheiden sich jedoch deutlich. Im einen Fall nutzen wir die Relation um in Teilmengen von K unsichere Elemente zu identifizieren, die dann gewissermaßen erst in einem zweiten Schritt entfernt werden. Im anderen Fall definiert die Relation Kontraktionen unmittelbar.

Die Grundidee ist wieder recht einfach. Man stelle sich vor, man beobachte einen Probanden, der immer wieder vor Auswahlaufgaben in einem bestimmten Bereich gestellt wird. Unter der Annahme, daß die Auswahl nicht zufällig ist, werden wir mit der Zeit Hypothesen über eine Präferenzordnung des Probanden in diesem Bereich aufstellen. Wir gehen davon aus, daß eine solche Ordnung des Bereichs der Auswahl zugrundeliegt, so daß wir auf ihrer Basis alle Auswahlen rekonstruieren bzw. vorhersagen können. [18]

Wenn wir vor die Aufgabe gestellt sind, eine Konjunktion $A \wedge B$ (in einer Theorie K) aufzugeben, dann müssen wir ebenfalls eine Wahl treffen: Wir können A oder B oder beide aufgeben. Wenn wir uns dafür entscheiden, A aufzugeben und B zu behalten, dann zeigen wir damit unsere (strikte) Präferenz von B über A, d.h. $A \prec B$ (in K) in genau dem folgenden Sinne:

(CE) $\qquad A \prec_{(K)} B$ gdw $A \notin K - A \wedge B$ und $B \in K - A \wedge B$.

Wir haben hier aus Kontraktionen von K eine Relation epistemischer Verankerung in K gewonnen. Aus den Eigenschaften von Kontraktionen können wir zudem auf Eigenschaften der Relation $\prec$ schließen:

---

[17] In [250] wird die Beziehung zwischen Sicherheit und epistemischer Verankerung genauer unter die Lupe genommen.

[18] Hier dient dieses Szenario nur als Heuristik, um epistemische Verankerung als durch die Wahl von Kontraktionsoptionen enhüllte Präferenz zu deuten. Daß dahinter sehr viel mehr als eine bloß hilfreiche Analogie steckt, zeigt Rott [251].

- Wenn Kontraktionen (über K) die Bedingungen (C1–8) erfüllen, dann hat die Relation $\prec_{(K)}$ nach (CE) die Eigenschaften [E ...].

Nun sind wir hier nicht an *irgendeinem* Bündel von $\prec$-Eigenschaften interessiert, sondern an einem, welches in einem bestimmten Sinne vollständig ist. (In der oben beschriebenen Analogie suchen wir nach einer Präferenzrelation, die es erlaubt *alle* erfolgten und künftigen Auswahlen nachzuvollziehen.) Vollständig wäre [E ...] genau dann, wenn daraus alle guten Kontraktionen erzeugt werden können. Dazu benötigen wir noch eine Bedingung [EC], die aus Relationen epistemischer Verankerung Kontraktionen erzeugt. Vor diesem Hintergrund müssen wir [E ...] dann so bestimmen, daß wir folgende Aussage machen können:

- Wenn $\prec_{(K)}$ die Eigenschaften [E ...] hat, dann erfüllen Kontraktionen (über K) nach [EC] die Bedingungen (C1–8).

Wie sind also die Andeutungen [EC] und [E ...] zu füllen? Wir beginnen mit

(EC) $\quad B \in K - A$ gdw $B \in K$ und ($A \prec A \vee B$ oder $A = \top$).

Anders als (CE), erschließt sich sicherlich nicht unmittelbar, warum (EC) dem Zweck dient. Gärdenfors und Makinson [101] argumentieren so:

(Dem Argument sei vorausgeschickt: Die Bedingung (CE) ist bereits gesetzt und wir dürfen sie daher verwenden. Ferner soll (EC) *gute* Kontraktionen erzeugen. Darunter verstehen wir hier solche, die (C1–8) erfüllen. Von diesen Bedingungen dürfen wir also auch Gebrauch machen.)

$\Rightarrow$: Wenn (1) $B \in K - A$, dann auch $B \in K$ nach (C3) und (2) $A \vee B \in K - A$. Wir nehmen nun weiter an, daß (3) $A \neq \top$ und zeigen, daß $A \prec A \vee B$. Aus (3) folgt nach (C2), (4) $A \notin K - A$. Aber nach (C6) ist $K - A = K - A \wedge (A \vee B)$. Also aus (2) und (4): $A \vee B \in K - A \wedge (A \vee B)$ und $A \notin K - A \wedge (A \vee B)$, was nach (CE) bedeutet, daß $A \prec A \vee B$.
$\Leftarrow$: Angenommen (1) $B \in K$. Im Fall $A = \top$ folgt aus den C-Postulaten, daß $K - A = K$. Also $B \in K - A$, wie gewünscht. Im Fall (2) $A \prec A \vee B$ haben wir nach (CE), $A \vee B \in K - A \wedge (A \vee B)$ und also (3) $A \vee B \in K - A$ nach (C6). Aus (1) folgt nach (C5), $B \in K - A + A$, also (4) $\neg A \vee B \in K - A$. Aber dann aus (3) und (4) wiederum: $B \in K - A$. ∎

Nun zu den Eigenschaften [E ...] epistemischer Verankerung. Wir werden gleich sehen, daß die folgenden Bedingungen geeignet sind:

($\prec$Irr) $\quad A \not\prec A$

($\prec$Auf) Wenn $A \prec B$ und $B \vdash C$, dann $A \prec C$.

($\prec$Ab) Wenn $A \vdash B$ und $B \prec C$, dann $A \prec C$.

($\prec$KAuf). Wenn $A \prec B$ und $A \prec C$, dann $A \prec B \wedge C$

($\prec$KAb). Wenn $A \wedge B \prec B$, dann $A \prec B$

Die letzten vier Bedingungen implizieren, daß $\prec$ transitiv ist:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A \prec B}{A \wedge B \prec B}\,\text{Ab} \qquad \dfrac{B \prec C}{A \wedge B \prec C}\,\text{Ab}}{A \wedge B \prec B \wedge C}\,\text{KAuf}}{A \wedge (B \wedge C) \prec B \wedge C}\,\text{Ab}}{A \prec B \wedge C}\,\text{KAb}}{A \prec C}\,\text{Auf}$$

Transitivität zusammen mit Irreflexivität schließt Zyklen der Form $A_0 \prec A_1 \cdots \prec A_k$ mit $A_0 = A_k$ aus. Jede Relation epistemischer Verankerung ist also eine Sicherheitsrelation, die zudem (Auf) und (Ab) erfüllt. Das Umgekehrte gilt natürlich nicht, da Sicherheitsrelationen auch unter (Auf) und (Ab) nicht allgemein die beiden Konjunktionsbedingungen erfüllen.

Wir wollen die Gesamtheit der oben genannten Bedingungen die *E-Bedingungen* nennen. Sie passen zu der intuitiven Interpretation von $A \prec B$ als "A ist einfacher aufzugeben als B". Betrachten wir beispielsweise die Konjunktionsbedingungen. (KAuf): Wenn A einfacher als B und als C aufzugeben ist, dann können wir nicht $B \wedge C$ aufgeben und A behalten. Denn dann müßten wir ja auch B oder C aufgeben und also – entgegen unserer Annahme – geneigt sein, eher B oder C als A aufzugeben. (KAb): Wenn $A \wedge B$ einfacher als B aufzugeben ist, dann muß das an A liegen, d.h. A muß einfacher als B aufzugeben sein.

Wichtiger als intuitive Stimmigkeit der E-Bedingungen ist jedoch das folgende Resultat.

Satz 18. (Gärdenfors und Makinson [101].)

1. *Wenn Kontraktionen (über K) die Bedingungen (C1–8) erfüllen, dann erfüllt die Relation $\prec_K$ nach (CE) die E-Bedingungen, sowie (EC).*
2. *Wenn $\prec_K$ die E-Bedingungen erfüllt, dann erfüllen Kontraktionen (über K) nach (EC) die Bedingungen (C1–8), sowie (CE).*

Beweis. Siehe [101].

Die Zusätze "sowie (EC)" und "sowie (CE)" zeigen eine besonders starke Beziehung zwischen Kontraktionen und Relationen epistemischer Verankerung an. Gälten die Zusätze nicht, dann wäre folgendes möglich. Wenn wir erst aus einer (C1–8)-Kontraktion $K -_1 (\ )$ per (CE) eine bestimmte E-Relation $\prec_1$ für K generieren, und dann aus genau dieser Relation eine (C1–8)-Kontraktion $K -_2 (\ )$ erzeugen, dann bliebe die Frage offen, ob $-_1 = -_2$ ist; d.h., ist für alle Formeln A, $K -_1 A = K -_2 A$? Ebenso in der umgekehrten Richtung.

Nach Satz 18 jedoch verhalten sich (CE) und (EC) ähnlich wie die Levi- und Harper-Identitäten zueinander; vgl. die Diskussion auf p. 499. Sei $-$ eine (C1–8)-Kontraktion und $e(-)$ die durch (CE) definierte E-Relation. Ebenso sei $\prec$ eine E-Relation und $c(\prec)$ die durch (EC) definierte (C1–8)-Kontraktion. Dann folgt aus dem Satz:

$$c(e(-)) = - \quad \text{und} \quad e(c(\prec)) = \prec .$$

Wir wollen $- : K \times \text{FML} \longrightarrow \wp(\text{FML})$ eine *E-Kontraktion* über K nennen, wenn es eine E-Relation $\prec_K$ gibt so, daß für jede Aussage A, $K - A$ die Bedingung (EC) erfüllt ist. Wir können nun unsere bisherigen Ergebnisse so zusammenfassen.

Für jede Operation $- : K \times \text{FML} \longrightarrow \wp(\text{FML})$ sind die folgenden Aussagen äquivalent:

- Die Operation erfüllt die Bedingungen (C1–8).
- Die Operation ist eine relationale Teilschnittkontraktion über K.
- Die Operation ist eine sichere Kontraktion über K unter den Bedingungen (VKon), und (Auf) oder (Ab).
- Die Operation ist eine E-Kontraktion.

In diesem Abschnitt standen – wie schon im vorigen – Kontraktionen im Mittelpunkt, da die Konzeption einer sicheren bzw. E-Kontraktion einfacher zu fassen ist. Über die Levi-Identität (LI) ergeben sich unmittelbar die entsprechenden Definitionen von Revisionen, die auf Sicherheits- bzw. E-Relationen beruhen. Die entsprechenden Repräsentationsresultate für Revisionen folgen als einfache Korollare. Im nächsten Abschnitt wird es einfacher sein, Revisionen in den Mittelpunkt treten zu lassen. Die Ableitung der entsprechenden Konstruktion von Kontraktionen geschieht dann über die Harper-Identität (HI).

Sphärensysteme. Der Ramsey-Test,

$$A \sqsupset B \in K \text{ gdw } B \in K * A$$

– obgleich er in dieser "naiven" Form den AGM-Postulaten nicht hinzugefügt werden darf –, deutet eine enge Beziehung an zwischen der Theorie der Revisionen und der Semantik (kontrafaktischer) Konditionale. Diese enge Beziehung lädt dazu ein, eine Theorie der Revisionen (und also auch der Kontraktionen) aus bekannten semantischen Theorien kontrafaktischer Konditionale herzuleiten. So erscheint der RT als epistemische Version eines Tests für die Wahrheit eines solchen Konditionals in einer Welt:

$A \sqsupset B$ ist in einer Welt a genau dann wahr, wenn B in $(a{*}A)$ wahr ist.

Semantische Theorien der Konditionale bestehen im Kern in einer Bestimmung des Ausdrucks $a{*}A$.[19] Gegeben eine solche Bestimmung von $a{*}A$ können wir uns fragen, ob diese eine bestimmte natürliche Verallgemeinerung zuläßt. Um welche Art der Verallgemeinerung es sich handeln muß, wird schnell deutlich, wenn wir uns daran erinnern, daß wir Überzeugungszustände als Mengen von möglichen Welten darstellen können: genau diejenigen Welten (Möglichkeiten), von denen das jeweilige Subjekt nicht ausschließen kann, daß es sich darin befindet. Wir fragen also danach, ob wir eine gegebene Bestimmung von $a{*}A$ auf natürliche Weise "zur Potenz erheben" können, also zu $K * A$ mit K einer Teilmenge der Menge aller möglichen Welten.

In der Semantik kontrafaktischer Konditionalsätze (siehe Kap. IV.3–5) gibt es viele äquivalente Modellierungen; siehe [182]. Zwei davon genießen gewissermaßen Standard-Status. Die eine macht Gebrauch von Auswahlfunktionen und bestimmt $a{*}A$ so:

$a{*}A$ ist der Schnitt aller in a ausgewählten A-Welten.

Teilschnittrevisionen lassen sich offensichtlich als Verallgemeinerungen dieses Ansatzes verstehen, Die andere Modellierung ordnet Welten um eine gegebene Welt in Sphären der Ähnlichkeit ("Nähe") an. In einem Sphärensystem wird $a{*}A$ so bestimmt:

$a{*}A$ ist die Menge derjenigen A-Welten, die sich in der a nächsten Sphäre, die A-Welten enthält, befinden (kurz: die Menge der a nächstgelegenen A-Welten).

[19] Der Ausdruck "$a{*}A$" hat hier lediglich eine heuristische Funktion. Die Revision (Änderung) einer Welt ist schon syntaktisch eine andere Operation als die Revision einer Theorie. Dennoch besteht eine starke Analogie zwischen Weltenrevision und Theorienrevision, deren systematischer Charakter deutlich wird, sobald wir Welten als maximal konsistente Theorien auffassen. In diesem Fall ist die obige Wahrheitsbedingung für Konditionale ein Spezialfall des Ramsey-Tests.

In Kapitel IV.5 haben wir Lewis' Sphärensemantik für kontrafaktische Konditionale dargestellt. Hier wollen wir die von Adam Grove [120] gefundene Überführung der Sphärensemantik in eine Theorie der Revisionen vorführen.

Wir gehen von einer Interpretation I aus, die den Atomen einer wahrheitsfunktionalen Sprache in Welten Wahrheitswerte zuweist. Die Grundmenge der Welten bezeichnen wir mit W. Die Erweiterung der Interpretationsfunktion für beliebige Formeln, $[\![\ ]\!] : \mathrm{FML} \longrightarrow \wp(W)$, definieren wir mittels der üblichen Rekursion: Für Atome P sei $[\![p]\!] = \{a \in W : I(P, a) = 1\}$; sodann

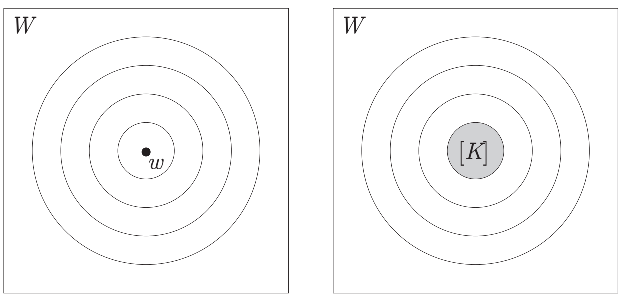
$$[\![\neg A]\!] = W \setminus [\![A]\!] \ (= \overline{[\![A]\!]}), \quad [\![A \wedge B]\!] = [\![A]\!] \cap [\![B]\!] \quad \text{usw.}$$

Diese Klammernotation dehnen wir so auf Mengen von Formeln aus:

$$[\![K]\!] = \{a \in W : \forall A \in K,\ a \in [\![A]\!]\}.$$

Statt $a \in [\![A]\!]$ schreiben wir gelegentlich, wie schon in früheren Kapiteln, $a \models A$. Man beachte, daß wenn $[\![X]\!] = [\![Y]\!]$, dann $\mathrm{Cn}(X) = \mathrm{Cn}(Y)$. Die Anwendung der Klammerfunktion auf Theorien ist also injektiv: Aus $[\![X]\!] = [\![Y]\!]$ folgt $X = Y$.

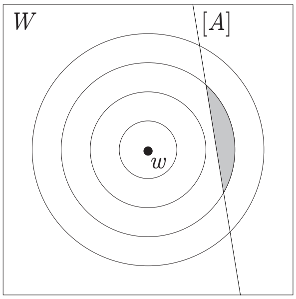
Im Gegensatz zu den Lewis'schen Sphären, haben die Grove'schen Sphären jetzt keine Welten zum Mittelpunkt, sondern Mengen von Welten, d.h. Objekte, die sich zur Modellierung von Überzeugungszuständen eignen. Die Sphären um einen Zustand K stellen eine Anordnung weiterer Mengen von Möglichkeiten dar (Rückfallmöglichkeiten), die das Subjekt im Zustand K zwar ausschließt, gegebenenfalls aber in dieser Reihenfolge in Betracht ziehen würde. Das folgende Diagramm stellt diese Idee dar. (Links ein Sphärensystem nach Lewis, rechts eines nach Grove.)

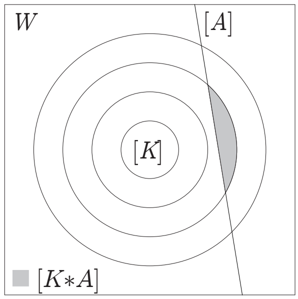


Das nächste linke Diagramm illustriert Lewis' Gebrauch der Sphären. Die Welt a liegt außerhalb von $[\![A]\!]$. Wenn nun kontrafaktisch (d.h. "kontra-a") die Möglichkeit A ins Auge gefaßt werden soll, dann sollten wir diejenigen

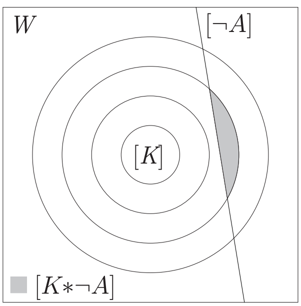
A-Welten betrachten, die nach der Sphärenordnung a am nächsten sind. Das ist der schraffierte Bereich im Diagramm. (Das kontrafaktische Konditional $A \sqsupset B$ ist nun in a genau dann wahr, wenn der schraffierte Bereich vollständig in den Bereich der B-Welten fällt.)

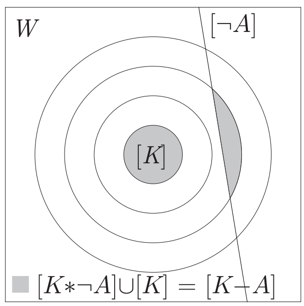
Das rechte Diagramm zeigt eine Sphärenanordnung um eine Theorie K, die in den Bereich $[\![\neg A]\!]$ fällt, also inkonsistent mit A ist. Vor die Aufgabe gestellt, K um A zu revidieren, sollten wir nun die naheliegendsten Rückfallmöglichkeiten aufsuchen, in denen A der Fall ist. Wie weit wir uns dazu von K entfernen müssen, wird durch das Sphärensystem um K angegeben. Im Diagramm stellt der schraffierte Bereich die neue Theorie $K * A$ dar.





Davon ausgehend lassen sich über die Harper-Identität $K - A = K * \neg A \cap K$ auch Kontraktionen darstellen:





Das Diagramm zeigt den Hauptfall $A \in K$, d.h. $[\![K]\!] \subseteq [\![A]\!]$. Falls $A \notin K$ – d.h. $[\![K]\!] \cap \overline{[\![A]\!]} \neq \emptyset$ –, dann ist $[\![K]\!]$ die kleinste Sphäre, die $\neg A$ erlaubt. Das ist $[\![K]\!]$ selbst; daher $[\![K - A]\!] = [\![K]\!]$. Das linke Diagramm zeigt, wie erklärt, die Revision von K um $\neg A$. Nun sind wir in einem Bereich, der für A mehr als offen ist: Er erzwingt $\neg A$. Das ist zu stark. Daher schwächen wir ab, indem wir $[\![K]\!]$ hinzunehmen, denn $[\![K * \neg A]\!] \cup [\![K]\!] = [\![K * \neg A \cap K]\!]$.

Wir wollen diese Ideen nun in Definitionen fassen, so daß ein neues Repräsentationsresultat möglich wird. Es sei W eine nichtleere Menge (von Möglichkeiten, möglichen Welten).

Definition 19. Ein *Sphärensystem* (in einer nichtleeren Grundmenge W) um eine Theorie K ist eine Mengenfamilie $\mathfrak{S}_K \subseteq \wp(W)$, so daß

1. $[\![K]\!] \in \mathfrak{S}_K$ und $\forall S \in \mathfrak{S}_K : [\![K]\!] \subseteq S$;
2. $W \in \mathfrak{S}_K$;
3. $\mathfrak{S}_K$ ist vollständig geordnet unter $\subseteq$ (d.h. $\forall S, S' \in \mathfrak{S}_K$: entweder $S \subseteq S'$ oder $S' \subseteq S$);
4. $\forall A \in$ FML: wenn $[\![A]\!] \neq \emptyset$, dann gibt es eine kleinste Sphäre $S \in \mathfrak{S}_K$ mit $[\![A]\!] \cap S \neq \emptyset$.

Bedingungen 1–3 ordnen Sphären um K wie eine Zwiebel (oder russische Puppe),[20] mit dem Kern $[\![K]\!]$ im Zentrum und W als äußerster Sphäre. Bedingung 4, bekannt als Limes-Annahme aus der Lewis'schen Sphärensemantik, schließt aus, daß $\mathfrak{S}_K$ (teilweise oder ganz) dicht geordnet ist, in welchem Fall es nicht immer eine kleinste Sphäre, die eine bestimmte Bedingung erfüllt, gibt; vgl. die Diskussion der Limes-Annahme in Kapitel IV.5.

Die Diagramme oben zeigten immer nur den Fall $[\![A]\!] \neq \emptyset$. In die Bestimmung von $K * A$ in einem Sphärensystem wollen wir jetzt auch den Fall $[\![A]\!] = \emptyset$ (d.h. A ist eine Kontradiktion) einbeziehen.

Für jede Menge $M \subseteq W$ sei die Beschränkung von $\mathfrak{S}_K$ durch M die Familie aller Sphären in $\mathfrak{S}_K$, die M nichtleer schneiden, d.h.

$$\mathfrak{S}_K/M = \{S \in \mathfrak{S}_K : S \cap M \neq \emptyset\}.$$

Wenn $M \neq \emptyset$, dann gibt es eine kleinste Sphäre in $\mathfrak{S}_K/M$, die wir mit $\min(\mathfrak{S}_K/M)$ bezeichnen wollen. Wenn $M = \emptyset$, dann ist $\min(\mathfrak{S}_K/M)$ nicht definiert und kann daher nicht zur Definition von Revisionen dienen. Deshalb definieren wir wie folgt:

$$(\mathfrak{S}*) \qquad [\![K * A]\!] = \begin{cases} \min(\mathfrak{S}_K/[\![A]\!]) \cap [\![A]\!], & \text{falls } [\![A]\!] \neq \emptyset; \\ \emptyset & \text{anderenfalls.} \end{cases}$$
$$K * A = \{B : [\![K * A]\!] \subseteq [\![B]\!]\}$$

Dieser Umgang mit dem Grenzfall $[\![A]\!] = \emptyset$ garantiert sofort das Postulat (R5). Wenn $*$ für eine bestimmte Theorie K und beliebige Formeln so wie in $(\mathfrak{S}*)$ definiert ist, dann wollen wir $*$ eine *Sphärenrevision* für K nennen.

[20] Die treffende Bezeichnung "Zwiebel" stammt von Segerberg.

Wie im Diagramm oben schon angedeutet lassen sich *Sphärenkontraktionen* so definieren:

$$(\mathfrak{S}-) \qquad [\![K - A]\!] = ([\![K * \neg A]\!]) \cup [\![K]\!].$$

Der nächste Satz besagt, daß Revisionen, welche die Begingungen (R1–8) erfüllen, genau diejenigen sind, die als Sphärenrevisionen darstellbar sind.

SATZ 20. Repräsentation für (R1–8) (Grove [120]). *Eine Revisionsoperation für K erfüllt genau dann die Bedingungen (R1–8), wenn es ein Sphärensystem $\mathfrak{S}_K$ gibt, so daß für alle Formeln A, $(\mathfrak{S}*)$ erfüllt ist.*

BEWEISSKIZZE. Es sei $\mathfrak{W}$ die Menge aller maximal konsistenten Formelmengen. Wir schreiben $[A]$ für $\{X \in \mathfrak{W} : K \subseteq X\}$. Unter hier zutreffenden Annahmen gibt es eine Bijektion zwischen $\mathfrak{W}$ und W und also zwischen $[A]$ und $[\![A]\!]$.

Für die $\Rightarrow$-Richtung definieren wir mithilfe einer (R1–8)-Revisionsfunktion für eine beliebige Theorie K eine Mengenfamilie $\mathfrak{M}_K$, von der wir dann zeigen, daß sie (a) die Bedingungen für ein Sphärensystem um K, sowie (b) für alle Formeln A die Gleichung $K * A = \{B : [K * A] \subseteq [B]\}$ (mit $[K * A]$ wie in $(\mathfrak{S}*)$ für $[\![\;]\!]$) erfüllt. Sei nun $\mathfrak{M}_K \subseteq \wp(\mathfrak{W})$ so, daß für jede Menge M in $\mathfrak{M}_K$:

1. $M = [A]$; oder
2. $M = \mathfrak{M}_K$; oder
3. $M \subseteq \bigcup_B [K * B]$ und $\forall B$: wenn $M \cap [B] \neq \emptyset$, dann $[K * B] \subseteq M$.

Die ersten zwei Klauseln entsprechen direkt den Bedingungen 1 und 2 in Definition 19. Der Sinn der dritten Klausel erschließt sich erst im Beweis. Sie stellt sicher, daß $\mathfrak{M}_K$ durch $\subseteq$ vollständig geordnet und die Limes-Annahme (Bedingung 4 der Definition) erfüllt ist. Sodann können wir beweisen, daß $\mathfrak{M}_K$ die Desiderata (a) und (b) erfüllt.

In der anderen Richtung müssen wir die Eigenschaften (R1–8) für Sphärenrevisionen nachweisen. Ein vollständiger Beweis in beide Richtungen ist in [129, Theorem 3.74] angegeben. ∎

Die Modellierung von Kontraktionen mittels Teilschnitten einerseits und mittels Sphärensystemen andererseits sehen auf den ersten Blick sehr verschieden aus. Aber tatsächlich gibt es eine enge Beziehung zwischen beiden. Um diese Beziehung darzustellen, wollen wir für jede Menge X von Welten mit $|X|_I = \{A : \forall a \in X : a \models_I A\}$ die Menge der in allen Welten in X (unter der Interpretation I) wahren Formeln notieren. Nun kann es in Sphärenmodellen durchaus Welten a und b geben, die zwar numerisch distinkt aber qualitativ gleich sind, indem sie genau dieselben Formeln wahr machen. Wir gewinnen jedoch an Einfachheit und verlieren nichts wesentliches, wenn wir jetzt nur solche Sphärenmodelle betrachten, in denen solche Zwillingswelten ausgeschlossen sind, d.h. in denen diese Implikation für alle $a, b \in W$ gilt:

$$(*) \qquad \text{Wenn } |a|_I = |b|_I, \text{ dann } a = b.$$

Es sei $A \in K$ und $\nvdash A$. Dann können wir folgende 1:1-Entsprechung zwischen Elementen in $K \perp A$ und solchen in $[\![\neg A]\!]$ beobachten. (Den beliebig gewählten Index I denken wir uns dazu.)

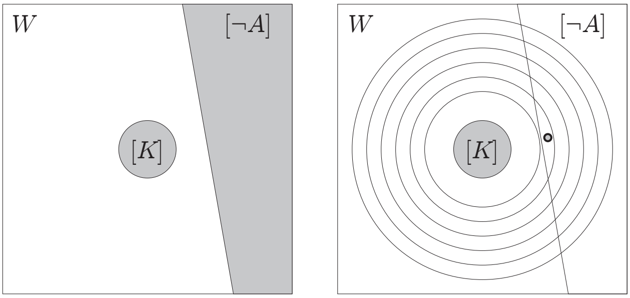
Beobachtung 21. (Grove [120].)

1. *Wenn $H \in K \perp A$, dann gibt es genau eine Welt in $[\![\neg A]\!]$, so daß $[\![H]\!] \cap [\![\neg A]\!] \neq \emptyset$.*
2. *Wenn $a \in [\![\neg A]\!]$, dann ist $|[\![K]\!] \cup \{a\}|$ in $K \perp A$.*

Beweis. *Ad* 1: Da $\nvdash A$, so ist $[\![\neg A]\!] \neq \emptyset$. Also folgt aus $H \in K \perp A$, daß die Menge $M = [\![H]\!] \cap [\![\neg A]\!]$ nicht leer ist. Wir zeigen, daß M eine Einermenge ist. Angenommen, $a, b \in M$ und $a \neq b$. Dann folgt aus $(*)$, daß es eine Formel B gibt, so daß (1) $a \notin [\![B]\!]$ und (2) $b \in [\![B]\!]$. Da $M \subseteq [\![\neg A]\!]$, so wissen wir auch, daß (3) $a, b \notin [\![A]\!]$. Aus (1) und (3), bzw. (2) und (3) erhalten wir $a \notin [\![A]\!] \cup [\![B]\!] = [\![A \vee B]\!]$ und $b \notin [\![A]\!] \cup [\![\neg B]\!] = [\![A \vee \neg B]\!]$. Da $a, b \in M \subseteq H$, so folgt daraus, daß (4) $[\![H]\!] \not\subseteq [\![A \vee B]\!]$ und $[\![H]\!] \not\subseteq [\![A \vee \neg B]\!]$. — Wir arbeiten jedoch unter der Annahme, daß $A \in K$, und also $A \vee B \in K$ und $A \vee \neg B \in K$. Da H maximal ist, so muß entweder $A \vee B$ oder $A \vee \neg B$ in H sein – im Widerspruch zu (4).

Ad 2: Sei $a \in [\![\neg A]\!]$. Wir zeigen, daß $H = |[\![K]\!] \cup \{a\}|$ in $K \perp A$ ist. Es ist klar, daß (a) $H \subseteq K$ und (b) $H \nvdash A$. Ebenso klar ist, daß für keine Menge $M \subset [\![K]\!] \cup \{a\}$, $|M|$ die Bedingungen (a) und (b) erfüllen kann. Also ist H maximal. ∎

Restemengen in $K \perp A$ korrespondieren also zu Weltenmengen $[\![K]\!] \cup \{a\}$, wobei $a \in [\![\neg A]\!]$. Im Falle von Kontraktionen nach dem *vollen Schnitt*, d.h. $K - A = \bigcap(K \perp A)$ bedeutet das, daß $[\![K - A]\!] = [\![K]\!] \cup [\![\neg A]\!]$, wie im folgenden Diagramm links zu sehen ist:



Für den vollen Schnitt betrachten wir sehr grobe Sphärensysteme um einen Überzeugungszustand K, nämlich solche mit nur zwei Sphären: der kleinsten Sphäre $\llbracket K \rrbracket$ und der größten Sphäre W. Da nach vollem Schnitt $\llbracket K - A \rrbracket = \llbracket K \rrbracket \cup \llbracket \neg A \rrbracket$, so ist $K - A = K \cap \mathrm{Cn}(\neg A)$. Das ist die Gleichung (C†) auf p. 510, deren Gültigkeit in der Sphärenmodellierung eine unmittelbare Folge der gerade beschriebenen Eins-zu-Eins-Korrespondenz ist.

Für Kontraktionen nach *Maximalwahl* wählen wir genau ein Element in $K \bot A$ aus, d.h. $K - A = H$, wobei $H \in K \bot A$. In Sphärensysteme übersetzt, bedeutet das, daß $\llbracket K - A \rrbracket = \llbracket K \rrbracket \cup \{a\}$ mit $a \in \llbracket \neg A \rrbracket$. Maximalwahl setzt also voraus, daß die Sphären um K so fein sind, daß die kleinste Sphäre, die eine (nichtleere) Proposition $\llbracket \neg A \rrbracket$ schneidet, im Schnitt nur genau ein Element enthält; vgl. das Diagramm oben rechts. (In der Semantik kontrafaktischer Konditionalsätze entspricht das der sogenannten Stalnaker-Annahme; siehe Kapitel IV.4–5.)

Wenn $A \in K$, dann enthält jede Menge $K \cap |a|$ mit $a \in \llbracket \neg A \rrbracket$ entweder $A \vee B$ oder $A \vee \neg B$, für beliebige Formeln B (vgl. den Beweis der Beobachtung oben). Da Kontraktionen nach Maximalwahl genau solche Mengen sind, wird auch in der Sphärenmodellierung unmittelbar deutlich, warum die Eigenschaft (C‡) (vgl. p. 510) für Maximalwahl charakteristisch ist.

Kontraktionen nach *Teilschnitt* liegen in der Mitte zwischen den beiden extremen Optionen: Das Sphärensystem wird nicht auf eine zu grobe Struktur reduziert, wie bei vollen Schnitten, und es wird nicht zu filigran, wie bei Maximalwahlen. Kontraktionen und Revisionen in Sphärensystemen nach Def. 19 erweisen sich als äquivalent zum Resultat von Teilschnitten von Restemengen.

8. Überzeugungswandel und anfechtbares Schließen

Die Teilschnittkonstruktionen streichen die Verwandtschaft der Theorie der Überzeugungsänderung mit Theorien anfechtbaren Schließens in besonders einfacher Weise heraus. Deshalb werden wir diese Beziehung gleich hier behandeln, bevor im folgenden Abschnitt drei weitere Modellierungen von Überzeugungsänderungen vorgestellt werden.

Auf den ersten Blick haben Theorien des Überzeugungswandels und Theorien des anfechtbaren (nicht-monotonen) Schließens sehr verschiedene Phänomene zum jeweiligen Gegenstand. Diese fragt nach im Prinzip riskanten aber normalerweise vernünftigen Schlüssen, die wir aus einer Menge von Annahmen ziehen können. Jene geht aus von einer Menge von Annahmen, die wir uns zu eigen gemacht haben, und fragt, wie wir diese vernünftigerweise ändern können, falls nötig. Bei genauerer Betrachtung wird jedoch eine enge Verbindung sichtbar. Anfechtbare Schlüsse sind anfechtbar, weil sie auf Hintergrundannahmen beruhen, deren Wahrheit nicht garantiert ist, obgleich sie normalerweise als wahr unterstellt werden dürfen. Anfechtbare Schlüsse sind nicht generell monoton, weil die explizit vorgegebenen Prämissen Hintergrundannahmen widersprechen können; diese müssen dann ausgeblendet werden. Je stärker die Prämissen sind, desto mehr Hintergrundannahmen müssen gegebenenfalls deaktiviert werden. Das kann dazu führen, daß eine Konklusion nach Verstärkung der Prämissen nicht mehr folgt. Die Deaktivierung von Hintergrundannahmen durch Prämissen können wir als eine *Revision* der Hintergrundannahmen betrachten. Denn statt zu sagen,

> B folgt aus der Prämisse A vor dem Hintergrund der Annahmen K,

können wir offenbar ebensogut sagen,

> B folgt aus dem Hintergrund K, nachdem dieser um die Prämisse A revidiert wurde.

Umgekehrt können wir die Beobachtung, daß B in der Revision von K nach A enthalten ist, auch als die Beobachtung eines anfechtbaren Schlusses betrachten, nämlich des Schlusses von A auf B vor dem Hintergrund der Annahmenmenge K. So gesehen sind anfechtbares Schließen und Überzeugungsrevision nur zwei Seiten derselben Medaille und die jeweiligen Theorien sollten ohne weiteres ineinander übersetzbar sein.

Nun haben wir in Kapitel VII gesehen, daß anfechtbares Schließen kein einheitliches Phänomen ist. Wenn wir dort zum Beispiel das Schließen unter der *Closed World Assumption* dazu gerechnet haben, dann ist dieses Schließen sicher in einem etwas anderen Sinne anfechtbar, als es das Schliessen ist, das von Reiters Default Logik ins Auge gefaßt wird. Diese, und viele andere Schlußrelationen sind nicht monoton—aber mit einer jeweils eigenen Erklärung dieser Eigenschaft. Ebenso ist die klassische AGM-Theorie Ausgangspunkt für viele Variationen, die jeweils besondere Arten des Überzeugungswandels zum Gegenstand haben. [21] Die Erwartung der gegenseitigen Übersetzbarkeit zwischen Theorien anfechtbaren Schließens einerseits und Theorien der Überzeugungsrevision andererseits über das gesamte Feld solcher Theorien muß also etwas gedämpft werden. Dennoch lassen sich typische Vertreter der beiden Theoriengattungen aufgrund der wechelseitigen Übersetzbarkeit von Postulaten gut miteinander vergleichen. Die folgende Korrespondenz zwischen den AGM-Postulaten für Revisionen und wichtigen Bedingungen für Operationen und Relationen anfechtbaren Schließens wurde erstmals von Gärdenfors und Makinson in [206] aufgezeigt; siehe auch [102], [103, Kap. 6] sowie [200, Kap. 6].

Die oben angesprochene Übersetzungsregel sieht so aus:

$$(\rightleftharpoons) \qquad B \in K * A \ \rightleftharpoons\ B \in \mathrm{C}_K(A) \text{ bzw. } A \mid\!\sim_K B.$$

Hier soll C für eine Schlußoperation, bzw. $\mid\!\sim$ für eine Schlußrelation, jeweils basierend auf einer Formelmenge K, stehen. Schließen aus (endlichen) Formelmengen X können wir simulieren, indem wir der Operation C_K die Konjunktion aller Elemente in X als Argument geben. Die so verstandene Schlußoperation ist nicht monoton in dem Sinne, daß stärkere Prämissen nicht allgemein stärkere Konklusionen generieren, d.h.

$$\text{nicht: Wenn } A \vdash B, \text{ dann } \mathrm{C}_K(B) \subseteq \mathrm{C}_K(A).$$

Das liegt daran, daß Revisionen nicht so sind, daß $K * B \subseteq K * A$, falls $A \vdash B$. (Man betrachte beispielsweise $A = B \wedge C$ ($\neq \perp$) und $K = \mathrm{Cn}(\neg(B \wedge C))$. Da $\neg B \notin K$, ist einerseits $K + B \subseteq K * B$ nach (R4). Also ist $\neg C \in K * B$. Andererseits ist $C \in K * B \wedge C$ ($= K * A$). Da nach $A \neq \perp$ und (R5) $K * A$ konsistent ist, kann $\neg C$ nicht in $K * A$ sein .)

Im Vorübergehen betrachten wir kurz eine alternative Weise, aus Revisionen nicht-monotone Operationen herzustellen. In der Übersetzungsregel oben tritt K in die Rolle der fix gehaltenen Hintergrundannahmen. In der alternativen Übersetzung $B \in K * A \rightleftharpoons\ B \in \mathrm{C}_A(K)$ übernimmt das Revisionsziel A diese Rolle und K repräsentiert die variierenden Prämissen für C_A. Solche Operationen C_A sind nicht monoton, da Revisionen es im oben schon besprochen Sinne (RM) nicht sind, d.h.

$$\textit{nicht}\text{: Wenn } K \subseteq H, \text{ dann } K * A \subseteq H * A.$$

[21] Hansson [129] enthält eine Übersicht über einige dieser Variationen.

Wenn H logisch mindestens so stark ist wie K, dann ist damit nicht garantiert, daß $\mathrm{C}_A(K) \subseteq \mathrm{C}_A(H)$. So weit, so gut. Jedoch erfüllt C_A – anders als C_K – nicht die für Schlußoperationen elementare Bedingung der Inklusion, $K \subseteq \mathrm{C}_A(K)$. (Im Hauptfall $\neg A \in K$ und $\not\vdash \neg A$ ist nie $K \subseteq K{*}A$.) So eignet sich diese Operation nicht für den angestrebten Vergleich mit anfechtbarem Schließen. Wir werden jedoch am Ende des Kapitels Gelegenheit haben, darauf zurückzukommen.

Wir wollen nun die Revisionspostulate in solche für C_K übersetzen. Da in den einzelnen Postulaten die Menge K immer konstant bleibt, lassen wir den Index K einfach fort; er ist aus dem Kontext immer eindeutig ergänzbar. Dann resultieren die folgenden Bedingungen für jede Operation C (basierend auf einer Menge K):

(RC1) $\mathrm{Cn}(\mathrm{C}(A)) \subseteq \mathrm{C}(A)$.

(RC2) $A \in \mathrm{C}(A)$.

(RC3) $\mathrm{C}(A) \subseteq \mathrm{Cn}(K \cup \{A\})$.

(RC4) Wenn $\neg A \notin \mathrm{Cn}(K)$, dann $\mathrm{Cn}(K \cup \{A\}) \subseteq \mathrm{C}(A)$.

(RC5) Wenn $\bot \in \mathrm{C}(A)$, dann $\vdash \neg A$.

(RC6) Wenn $A \equiv B$, dann $\mathrm{C}(A) = \mathrm{C}(B)$.

(RC7) $\mathrm{C}(A \wedge B) \subseteq \mathrm{Cn}(\mathrm{C}(A) \cup \{B\})$.

(RC8) Wenn $\neg B \notin \mathrm{C}(A)$, dann $\mathrm{Cn}(\mathrm{C}(A) \cup \{B\}) \subseteq \mathrm{C}(A \wedge B)$.

Einige dieser Bedingungen sind alte Bekannte aus Kapitel VII; andere können wir aus solchen ableiten. Die RC-Bedingungen charakterisieren jedenfalls eine Familie von Operationen, die als Operationen (bzw. Relationen) anfechtbaren Schließens gelten dürfen. Wir wollen kurz einige der RC-Bedingungen, zur Abwechslung in relationalem Gewand, besprechen. (Dabei buchstabieren wir gegenüber den Formulierungen oben teilweise um.)

Die Regeln

$$\text{(RC1)}\ \frac{A \mid\!\sim B \quad B \vdash C}{A \mid\!\sim C} \qquad \text{(RC2)}\ A \mid\!\sim A \qquad \text{(RC6)}\ \frac{A \equiv B \quad A \mid\!\sim C}{B \mid\!\sim C}$$

kennen wir bereits aus Kapitel VII unter den Namen *Rechte Abschwächung* (RW), *Reflexivität* bzw. *Linke Logische Äquivalenz* (LLE). Das sind sehr elementare Eigenschaften des Schließens, die durch die Möglichkeit der Anfechtbarkeit nicht berührt werden. Die Regel der *Konsistenzerhaltung*,

$$\text{(RC5)}\ \frac{A \mid\!\sim \bot}{A \vdash \bot},$$

besagt, daß anfechtbares Schließen konsistent ist, es sei denn die Prämissen des Schließens sind (deduktiv) inkonsistent. Das ist eine sehr plausible Bedingung, welche $\mid\!\sim$-Relationen normalerweise erfüllen. [22]

$$\text{(RC3)}\ \frac{A \mid\!\sim C}{\top \mid\!\sim A \to C} \quad \text{und} \quad \text{(RC4)}\ \frac{\top \not\mid\!\sim \neg A \quad \top \mid\!\sim A \to C}{A \mid\!\sim C}$$

fallen unter die allgemeineren Bedingungen

$$\text{(RC7)}\ \frac{A \wedge B \mid\!\sim C}{B \mid\!\sim A \to C} \quad \text{und} \quad \text{(RC8)}\ \frac{B \not\mid\!\sim \neg A \quad B \mid\!\sim A \to C}{A \wedge B \mid\!\sim C}.$$

Es handelt sich um Ein- und Ausführungsregeln für die Implikation im Konsequens, wie sie auch für $\vdash$ gelten. Auch in diesem Fall berührt die Anfechtbarkeit von $\mid\!\sim$ nicht zwingend die Zulässigkeit des Schließens nach diesen Regeln. In dieser Form treten die Regeln in der Literatur über nicht-monotones Schließen selten auf. Jedoch ist (R7) im Kontext der übrigen Revisionspostulate äquivalent zu

$$\text{(R7}'\text{)} \qquad K * A \cap K * B \subseteq K * (A \vee B),$$

dessen Übersetzung

$$\text{(OEA)} \qquad \frac{A \mid\!\sim C \quad B \mid\!\sim C}{A \vee B \mid\!\sim C}$$

nichts anderes ist als die aus Kapitel VII bekannte *Oder-Einführung im Antezedens*. (RC8) impliziert das Prinzip der *Rationalen Monotonie*,

$$\text{(RatM)} \qquad \frac{B \not\mid\!\sim \neg A \quad B \mid\!\sim C}{A \wedge B \mid\!\sim C}.$$

Dies ist ein relativ starkes Prinzip, d.h. Monotonie unter einer relativ schwachen Einschränkung. Es kennzeichnet eine Familie nicht-monotoner Schlußrelationen, die in einer natürlichen Klasse sogenannter präferenzieller Modelle gelten. [23]

Wie steht die AGM-Theorie zu den folgenden drei Regeln, von denen gewöhnlich erwartet wird, daß sie zur Grundausstattung einer Logik anfechtbaren Schließens gehören (siehe Kapitel VII)?:

$$\text{(CT)}\ \frac{A \mid\!\sim B \quad A \wedge B \mid\!\sim C}{A \mid\!\sim C} \qquad \text{(CM)}\ \frac{A \mid\!\sim B \quad A \mid\!\sim C}{A \wedge B \mid\!\sim C} \qquad \text{(K+)}\ \frac{A \vdash B}{A \mid\!\sim B}.$$

[22] Eine Ausnahme aus "technischen" Gründen wird in [200, p. 71] erwähnt.

[23] Vgl. auch die ausführlichere Diskussion dieser und anderer Bedingungen für nicht-monotones Schließen im Vergleich zu Revisionspostulaten in [206], [102] und [103, Kap. 6].

Vorsichtige Monotonie (CM) erlaubt, wie (RatM), Prämissenerweiterung unter einer Bedingung; *kumulative Transitivität* (CT) ist eine eingeschränkte Schnittregel; *Supraklassität* (K+) reflektiert die leitende Idee, daß anfechtbares Schließen das deduktive Schließen erweitert. Wenn wir diese Regeln in Revisionsprinzipien übersetzen, so erhalten wir:

(RCT) Wenn $B \in K * A$, dann $K * A \wedge B \subseteq K * A$.

(RCM) Wenn $B \in K * A$, dann $K * A \subseteq K * A \wedge B$.

(RK+) Wenn $A \vdash B$, dann $B \in K * A$.

Das letzte Prinzip folgt aus den Bedingungen (R2), daß Revisionen erfolgreich, und (R1), daß Revisionen deduktiv abgeschlossen sein sollen. (RCT) und (RCM) folgen jeweils unmittelbar aus (R7) und (R8).

Nach diesem kurzen Vergleich von Postulaten für Revisionen mit solchen für anfechtbares Schließen, wollen wir uns nun ansehen, welche Rolle Teilschnittkonstruktionen in der Modellierung anfechtbaren Schließens spielen.

Auf p. 515 haben wir Revisionen nach Teilschnitt so definiert:

(1) $$K * A = \mathrm{Cn}(\bigcap s(K \perp \neg A) \cup \{A\}).$$

Nach der Übersetzung ($\rightleftharpoons$) erhalten wir aus (1) eine Darstellung der aus Revisionen gewonnenen Schlußoperation C^* mithilfe von Teilschnitten:

(2) $$\mathrm{C}^*_K(A) = \mathrm{Cn}(\bigcap s(K \perp \neg A) \cup \{A\}).$$

In Kapitel VII haben wir Folgerungen aus Default-Annahmen nach Teilschnitt so definiert:

(3) $$\mathrm{C}_K(A) = \bigcap \{\mathrm{Cn}(K' \cup \{A\}) : K' \in s(K \perp \neg A)\}.$$

(In Kap. VII haben wir die Menge $K \perp \neg A$ mit $K_{\{A\}}$ notiert.) Die Reihenfolge der Operation ist vertauscht: In (2) werden erst die ausgewählten Elemente geschnitten, danach wird A hinzugefügt und dann folgt der Abschluß unter Cn (kurz: $\bigcap, \cup, \mathrm{Cn}$); in (3) wird zu den ausgewählten Elementen A hinzugefügt, dann unter Cn abgeschlossen und danach geschnitten (kurz: $\cup, \mathrm{Cn}, \bigcap$). Es stellt sich die Frage, in welchem Verhältnis C^* und C zueinander stehen. Antwort: Wenn K eine Theorie ist, dann ist die Reihenfolge gleichgültig.

BEOBACHTUNG 22.
1. *Für jede Formel A und jede Menge K ist* $\mathrm{C}^*_K(A) \subseteq \mathrm{C}_K(A)$.
2. *Für jede Formel A und jede Theorie K ist* $\mathrm{C}_K(A) \subseteq \mathrm{C}^*_K(A)$.
3. *Wenn K eine Theorie ist, dann ist* $\mathrm{C}^*_K = \mathrm{C}_K$.

BEWEIS. Für 1 genügt es festzustellen, daß allgemein $\mathrm{Cn}(M \cap N) \subseteq \mathrm{Cn}(M) \cap \mathrm{Cn}(N)$. Denn, da $M \cap N \subseteq M$, folgt aus der Monotonie von Cn, daß $\mathrm{Cn}(M \cap N) \subseteq \mathrm{Cn}(M)$. Ebenso ist $\mathrm{Cn}(M \cap N) \subseteq \mathrm{Cn}(N)$ und also $\mathrm{Cn}(M \cap N) \subseteq \mathrm{Cn}(M) \cap \mathrm{Cn}(N)$.

Die Umkehrung in 2 gilt nur für Theorien. Denn in diesem Fall ist jedes Element einer Restefamilie eine Theorie, d.h. $(*)\ \forall K' \in K \bot \neg A : K' = \mathrm{Cn}(K')$. Ohne Einschränkung der Allgemeinheit betrachten wir nun den Fall $s(K \bot \neg A) = \{K', K''\}$. Dann

$$\begin{aligned} \mathrm{C}_K(A) &= \mathrm{Cn}(K' \cup \{A\}) \ \cap \ \mathrm{Cn}(K'' \cup \{A\}) \\ &\subseteq \mathrm{Cn}((\mathrm{Cn}(K') \cap \mathrm{Cn}(K'')) \cup \{A\}) \\ &= \mathrm{Cn}((K' \cap K'') \cup \{A\}) \qquad (*)! \\ &= \mathrm{C}^*_K(A). \end{aligned}$$

■

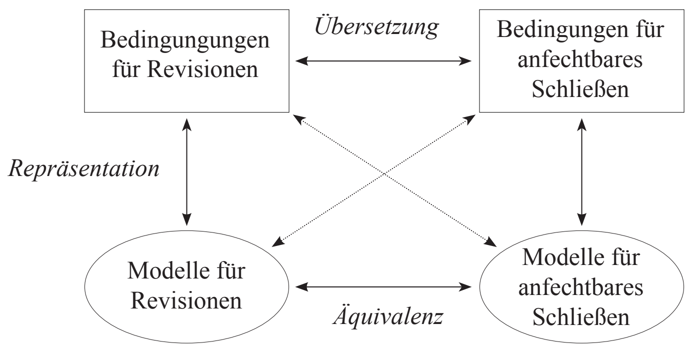
Die Bedingung in Teil 2 der Beobachtung, nämlich daß Hintergrundannahmen deduktiv abgeschlossen sein sollen, ist eine sehr natürliche. Sie ist äquivalent zu der folgenden, die wir ebenfalls schon in Kapitel VII besprochen haben:

$$\text{(DLE)} \qquad \frac{A \mid\!\sim_K B \quad H \equiv K}{A \mid\!\sim_H B} \qquad \text{bzw.} \qquad K \equiv H \ \Rightarrow\ C_K = C_H\,.$$

(DLE) besagt, daß es nur auf die logische Stärke der Hintergrundannahmen, nicht auf ihre syntaktische Kodierung ankommt. So soll zwischen den Hintergrundmengen $\{p \wedge q\}$ und $\{p, q\}$ nicht unterschieden werden, denn es kann kaum vernünftig sein, daß diese unter Umständen mehr Schlüsse zuläßt als jene.

Es gibt einen Grund, auf den Makinson in [200, p. 146] hinweist, warum die Reihenfolge des Schneidens und Abschließens unter Cn für C^* so ist wie in (2) – obgleich das unter der gerade besprochenen natürlichen Annahme letztlich keinen Unterschied macht. In der AGM-Theorie sind Kontraktionen die einfacheren Operationen, aus denen Revisionen mittels der (LI) entstehen, d.h. $K * A = \mathrm{Cn}(K - \neg A \cup \{A\})$. In Teilschnittkonstruktionen stellen sich Revisionen nach der (LI) daher als die rechte Seite von (2) dar. Warum ist die Reihenfolge für C in (3) anders? In der Theorie des anfechtbaren Schließen spielen Kontraktionen keine Rolle und können also auch keine Präferenz für die Gleichung (2) begründen. Für (3) spricht dann, daß im allgemeineren Fall, d.h. ohne die Bedingung (DLE), die in (3) definierte Operation stärker ist als die in (2). Selbst mit der Bedingung (DLE) in Kraft, mag für die Definition (3) angeführt werden, daß sie heuristisch naheliegender ist: Wir fragen danach, was aus allen mit der Prämisse A verträglichen Teilmengen der Hintergrundannahmen folgt, wenn wir A hinzunehmen – das ist die Reihenfolge $\cup, \mathrm{Cn}, \bigcap$.

Das folgende Diagramm (entlehnt aus [103, p. 106]) faßt zusammen, in welchem Sinne Theorienrevisionen und anfechtbares Schließen sich tatsächlich als zwei Seiten derselben Medaille erweisen:



Die senkrechten Pfeile deuten Repräsentationsresultate an: links für Revisionspostulate, rechts für allgemeine Bedingungen, unter welche anfechtbares Schließen gestellt werden kann. Der waagerechte Pfeil oben weist auf die Übersetzung ($\rightleftharpoons$) zwischen den beiden Postulatgruppen hin. Der syntaktischen Übersetzbarkeit entspricht semantische Äquivalenz, angedeutet durch den waagerechten Pfeil unten. Nun "kommutiert" das Diagramm, d.h. in den Diagonalen deuten sich neue Repräsentationsresultate an. Das ist Gegenstand der Beobachtung 22. Diese fußt auf den Repräsentationsresultaten 10 und 14 für die Revisionspostulate im Hinblick auf Teilschnittmodelle. In diesem Fall erklärt sich die Äquivalenz zwischen Modellen für Revisionen und Modellen für anfechtbares Schließen aus dem Umstand, daß für deduktiv abgeschlossene Mengen von Hintergrundannahmen die rechten, definierenden Seiten von (2) und (3) äquivalent sind.

9. Überzeugungswandel und Parakonsistenz

In der AGM-Theorie stellt sich die Revision eines Überzeugungssystems als Versuch dar, neue Information in dieses auf möglichst konservative Weise und zugleich widerspruchsfrei einzugliedern. Die Postulate und Konstruktionen sind so gewählt, daß beiden Forderungen Rechnung getragen wird. Revisionen sind also Operationen, die Information *konsistenzerhaltend* verarbeiten und dabei unter einer Maximierungsbedingung (im Hinblick auf den Wert schon vorhandene Information) stehen.

Eigentlich ist die Bedingung der Konsistenzerhaltung *per se* kein wesentliches Ziel. Durch die weitere Forderung, daß Überzeugungen unter klassischer logischer Folgerung abgeschlossen sein sollen, ist die Bedingung jedoch äquivalent zur Forderung, daß das Resultat einer Revision *nicht-trivial* sein, d.h. nicht alle Sätze der Sprache enthalten soll.[24] Gäbe es diese Äquivalenz nicht, könnten wir ein widersprüchliches Paar A und $\neg A$ genauso gut oder schlecht tolerieren wie z.B. das Paar “Dieses Feld ist völlig rot” (R) und “Dieses Feld ist völlig grün” (G). Im allgemeinen sind R und G sicher ein unerwünschtes Paar von Überzeugungen, denn nichts kann zugleich völlig rot und grün sein. Die AGM-Theorie zwingt uns jedoch nicht, die Logik um Axiome wie $R \leftrightarrow \neg G$ zu erweitern, seien diese semantischer, physikalischer oder anderer Natur. Der Unterschied zwischen $A \wedge \neg A$ und $R \wedge G$ besteht nicht darin, daß $R \wedge G$ relativ glaubhafter ist als $A \wedge \neg A$ (gleichgültig, wofür A steht). Er besteht vielmehr darin, daß $A \wedge \neg A$ Trivialität logisch erzwingt, während $R \wedge G$ das nicht tut, jedenfalls nicht ohne weiteres. Revisionen sind also Operationen, die Information *trivialitätsmeidend* und maximierend (im oben angedeuteten Sinne) verarbeiten. Das ist eine äquivalente aber den Zweck besser treffende Beschreibung.

So gesehen, teilt die AGM-Theorie der Revisionen das wesentliche Anliegen der meisten parakonsistenten Logiken. In beiden Unternehmungen geht es darum, aus möglicherweise inkonsistenter Information interessante Schlüsse zu ziehen. Im AGM-Fall treffen Information K und Information A aufeinander. Es findet dann eine Aufbereitung beider Informationen (“Prämissen”) statt, wobei der Input A *per* Annahme privilegiert ist (die Erfolgsbedingung R2).[25] Das Ergebnis der Aufbereitung ist so, daß daraus interessant klassisch geschlossen werden kann, auch wenn $K \cup \{A\}$ logisch

---

[24] Ausgenommen von der Forderung ist die Revision um einen Widerspruch – ein nicht weiter interessanter Fall, da wir Einladungen zu solchen Revisionen immer ablehnen sollten.

[25] Diese Annahme ist nicht zwingend für eine Theorie im AGM-Stil. Hanssons Theorie der nicht-priorisierten Überzeugungsänderung [128] klammert die Erfolgsbedingung R2 für Revisionen aus und schließt so die Möglichkeit ein, daß der Input A auf eine Theorie K trifft mit dem Resultat, daß A *nicht* übernommen wird. Ebenso, und etwas allgemeiner, verfährt die Verschmelzungsoperation (*merge*) in [85, 86].

widersprüchlich ist. Theorien parakonsistenten Schließens stehen typischerweise unter Bedingungen, die da Costa [56] so beschreibt:

> (i) In these calculi the principle of contradiction, $\neg(A \wedge \neg A)$, must not be a valid schema;
> (ii) from two contradictory formulas, P and $\neg P$, it will not in general be possible to deduce an arbitrary formula Q;
> (iii) it must be simple to extend [these calculi] to corresponding predicate calculi (with or without equality) of first order;
> (iv) [these calculi] must contain the most part of the schemata and rules of [classical logic] which do not interfere with the first conditions.
> (Evidently, the last two conditions are vague.)

Die zweite Bedingung versteht sich von selbst. Die erste Bedingung ist oft – aber nicht immer! – eine Voraussetzung für die zweite. Die dritte Bedingung wird für die beabsichtigten Anwendungen benötigt, z.B. für eine parakonsistente Arithmetik. Die vierte Bedingung soll sicherstellen, daß der Abschluß unter parakonsistenter Folgerung die Informativität von Prämissen nicht unplausibel schwächt. Insbesondere sollte aus widerspruchsfreien Prämissen parakonsistent genausoviel Information gewonnen werden können wie klassisch. Das ist in der Tat vage, denn es hängt davon ab, wie wir den Informationsgehalt von Satzmengen vergleichen. Aber die Absicht ist klar genug: Es geht darum, aus möglicherweise inkonsistenten Prämissen trivialitätsmeidend und zugleich informationsmaximierend zu schließen.

AGM-Revisionen und parakonsistente Folgerung verfolgen ein gleiches – oder doch ein sehr ähnliches – Ziel, tun das aber auf sehr unterschiedliche Weisen. Die AGM-Theorie verfährt in zwei Schritten: 1. Aufbereitung der Information unter Rückgriff auf eine Präferenzkomponente, 2. Extraktion der Information mittels klassischer Folgerung. Wir wollen uns das hier etwas näher ansehen, indem wir eine einfache AGM-repräsentierbare parakonsistente Folgerungsrelation betrachten.

Im Abschnitt über Überzeugungswandel und anfechtbares Schließen (p. 535) haben wir kurz diese Folgerungsoperation [bzw. Relation] betrachtet:

$$B \in C_A(K)\ [K \mid\!\sim_A B] \ \text{ gdw } \ B \in K * A.$$

Wir haben sie zunächst verworfen, da sie z.B. die Bedingung

$$\text{wenn } B \in K, \text{ dann } K \mid\!\sim_A B$$

nicht erfüllt. Jetzt wollen wir diese Relation doch etwas näher betrachten, und zwar für den Fall $A = \top$, d.h.

$$K \mid\!\sim_\top B \ \text{ gdw } \ B \in K * \top.$$

Mit der Levi- als auch mit der Harper-Identität ist übrigens $K * \top = K - \bot$ leicht nachweisbar. Hansson [125] nennt eine Kontraktion um $\bot$ eine *Konsolidierung*; die Verschmelzungsoperation (*merge*) in [85, 86] ist eine Verallgemeinerung. In der Standard AGM-Theorie ist K eine klassische Theorie. Aber wir können davon abstrahieren und K als beliebige Satzmenge auffassen (und schreiben dann X). Dazu müssen wir die AGM-Theorie so verallgemeinern, daß Revisionen und Kontraktionen auf beliebige Satzmengen (Basen) anwendbar sind. Die Postulate und Konstruktionen werden dabei etwas angepaßt; siehe die Bemerkungen auf p. 500. Ohne hier ins Detail zu gehen, können wir jedenfalls definieren (für beliebige Formelmengen X):

$$(\mid\!\sim_\top) \qquad X \mid\!\sim_\top A \text{ gdw } X * \top \vdash A.$$

Natürlich ist die Bedingung $X, A \mid\!\sim_\top A$ nach wie vor nicht erfüllt. Aber das ist nun bestens begründet, denn wenn $X \cup \{A\}$ inkonsistent ist, dann kann jede Formel in dieser Menge, dem Aufarbeitungsschritt, d.h. der Revision um $\top$, zum Oper fallen.

Sehen wir uns nun die da Costa-Bedingungen für parakonsistente Folgerung an:

(i) Diese Bedingung übersetzen wir hier so, daß es Mengen X geben muß, für die $X \mid\!\sim_\top \neg(A \wedge \neg A)$ falsch ist. Diese $\mid\!\sim_\top$-Sequenz gilt jedoch immer, da wir Information klassisch extrahieren und $\neg(A \wedge \neg A)$ ein Theorem der klassischen Logik ist, also aus beliebigen Mengen folgt.

(ii) Wir wollen nicht $X, A, \neg A \mid\!\sim_\top B$ für beliebige Wahlen von X, A und B. Die Bedingung ist offensichtlich erfüllt.

(iii) Die der Definition von $\mid\!\sim_\top$ zugrundeliegene AGM-Theorie ist umstandslos auf Sprachen erster Stufe anwendbar.

(iv) Wir können zeigen, daß diese Bedingung auf die stärkste mögliche Weise erfüllt ist. Denn wenn X konsistent ist (d.h. $X \nvdash \bot$), dann ist $X * \top = X$ und also koinzidieren $\mid\!\sim_\top$ und $\vdash$ in diesem Fall, d.h. $X \mid\!\sim_\top A$ gdw $X \vdash A$.

Bis auf (i) erfüllt $\mid\!\sim_\top$ also alle Bedingungen für gute parakonsistente Folgerung. Wie wichtig ist die erste Bedingung? Das hängt zum Teil davon ab, wie eng man parakonsistente Logik mit einer dialethischen Wahrheitstheorie verknüpft; siehe dazu [222, 225]. Einer solchen Theorie zufolge, gibt es *wahre* Instanzen von $A \wedge \neg A$ (sei A beispielsweise der Lügnersatz). Eine Logik mit $\neg(A \wedge \neg A)$ als Axiom (oder Theorem), würde diese Möglichkeit logisch ausschließen – jedenfalls unter jeder naheliegenden Interpretation der

Formel. Eine solche Logik könnte zwar parakonsistent im Buchstabensinne sein, wäre es aber aus dialethischer Perspektive dem Geiste nach nicht. Wenn man – wie z.B. da Costa – Parakonsistenz und Dialethismus nicht auf diese Weise verknüpft sieht, dann kann es technische Gründe geben, $\neg(A \wedge \neg A)$ als Theorem abzulehnen. So gilt für da Costas **C**-Systeme, daß sie in die klassische Logik zurückfallen, sobald man sie um das Axiomenschema $\neg(A \wedge \neg A)$ erweitert. So gesehen, ist die Bedingung (i) nur wichtig, wenn sie eine notwendige Bedingung für die entscheidende Bedingung (ii) darstellt. Leztere wird aber von $\mid\!\sim_{\top}$ auch ohne (i) erfüllt.

Wir haben also zwei Ansätze zu parakonsistentem Schließen vor uns. Der eine ist "nichttraditionell klassisch", da er auf der klassischen Folgerungsrelation beruht und diese mit einer vorgelagerten, durch einen Auswahlmechanismus gesteuerten Prämissenaufbereitung kombiniert. Der zweite Ansatz ist "traditionell nichtklassisch", da er Folgerung auf eine echte Teilmenge klassischer Folgerung beschränkt und dazu den traditionellen Weg über die Axiomatik oder Semantik wählt.

Das ist eine typische Situation in der modernen Logik. Schließen ist ein sehr vielfältiges Phänomen. Der traditionelle Zugang zu dieser Vielfalt bestand meist darin die Dominanz der klassischen Logik infrage zu stellen. Auf diese Weise wurden eine Reihe nichtklassischer Logiken vorgeschlagen. Auf dem Gebiet der Wissensrepräsentation – für das das Anliegen parakonsistenter Logik doch eigentlich einschlägig ist – haben sich die nichtklassischen Ansätze jedoch kaum durchsetzen können. Stattdessen zeigte sich, daß sich die Vielfalt des Schließens auch klassisch abbilden läßt: Die klassische Kerntheorie des Schließens wird dabei mit Modulen der Gewichtung und der Auswahl angereichert. (Vgl. dazu auch das Kapitel VII über anfechtbares Schließen.) Van Benthem [29] drückt die Erwartung aus, daß sich in manchen Fällen beide Ansätze zur Deckung bringen lassen. Jedoch sind die Ansätze – zugespitzt beschrieben: axiomatisch vs. prozedural – so verschieden, daß das Ausbleiben einschlägiger Repräsentationsresultate nicht überraschen kann.

Wir haben eingangs festgestellt, daß die AGM-Theorie der Revisionen und parakonsistente Theorien des Schließens das gemeinsame Anliegen haben, auch aus inkonsistenter Information interessante, und das heißt zumindest nicht-triviale Schlüsse zu ziehen. Was dieses Anliegen angeht, mag man mit guten Gründen dem nicht-traditionellen klassischen Ansatz von AGM den Vorzug geben. Aber parakonsistente Logiken können darüberhinaus einem Zweck dienen, der für AGM erst gar nicht zur Debatte steht: Eine parakonsistente Logik kann um Axiome zur Theorie eines spezifischen Bereichs erweitert werden. Die prominentesten Beispiele sind Theorien der Arithmetik und der Mengenlehre, insbesondere der naiven, klassisch inkonsistenten Mengenlehre. So können wir z.B. der klassischen Theorie **PA** ("Peano-Arithmetik") die Erweiterung einer parakonsistenten Logik (erster Stufe) um dieselben arithmetischen Axiome gegenüberstellen, zum Beispiel die Erweiterung der Relevanzlogik **R** zur arithmetischen Theorie $\mathbf{R}^{\#}$ (siehe Kap. VI). Wir können dann fragen, wieviel Arithmetik sich, im Vergleich mit **PA**, in $\mathbf{R}\#$ herleiten läßt. Im Falle von $\mathbf{PA} * \top$ ist diese Frage uninteressant, da sie eine triviale Antwort hat: $\mathbf{PA} * \top = \mathbf{PA}$. Im Falle der naiven Mengenlehre **NM** ist $\mathbf{NM} * \top$ kaum mehr als die Aufforderung eine konsistente Teiltheorie (z.B. **ZF**) zu wählen. Wie immer, sollte sich die Wahl der Mittel an den Zwecken ausrichten – eine Maxime, mit der jedes Kapitel dieses Buches und also das Buch insgesamt gut schließen kann.

LITERATUR

[1] Ackermann, W. Begründung einer strengen Implikation. *The Journal of Symbolic Logic 21*, 2 (1956), 113–128.

[2] Adams, E. *The Logic of Conditionals.* Reidel, Dordrecht, 1975.

[3] Adams, E. Four probability-preserving properties of inferences. *Journal of Philosophical Logic 25* (1996), 1–24.

[4] Adams, E. *A Primer of Probabilistic Logic.* CSLI Publications, Stanford, 1998.

[5] Alchourrón, C., Gärdenfors, P., und Makinson, D. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic 50*, 2 (1985), 510–530.

[6] Alchourrón, C., und Makinson, D. Hierarchies of regulations and their logic. In *New Studies in Deontic Logic*, R. Hilpinen, Hg., Reidel, Dordrecht, 1981, pp. 123–148.

[7] Alchourrón, C., und Makinson, D. On the logic of theory change: Contraction functions and their associated revision functions. *Theoria 48*, 1 (1982), 24–37.

[8] Alchourrón, C., and Makinson, D. Maps between some different kinds of contraction function: The finite case. *Studia Logica 45*, 2 (1984), 187–198.

[9] Alchourrón, C., und Makinson, D. On the logic of theory change: Safe contraction. *Studia Logica 44* (1985), 405–422.

[10] Anderson, A., und Belnap, N. *Entailment: The Logic of Relevance and Necessity, Vol. 1.* Princeton University Press, Princeton, 1975.

[11] Anderson, A., Belnap, N., und Dunn, J. *Entailment: The Logic of Relevance and Necessity, Vol. 2.* Princeton University Press, Princeton, 1992.

[12] Anderson, A. R., und Belnap, N. Tautological entailment. *Philosophical Studies 13*, 1-2 (1962), 9–24.

[13] Antoniou, G. A Tutorial on Default Logics. *ACM Computing Surveys 31*, 4 (1999), 337–359.

[14] Appiah, A. *Assertion and Conditionals.* Cambridge University Press, Cambridge, 1985.

[15] Arló-Costa, H., und Shapiro, S. Maps between conditional logic and non-monotonic logic. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference*, B. Nebel, C. Rich, und W. Swartout, Hgg., Morgan Kaufmann, San Mateo, Cal., 1992, pp. 553–565.

[16] Arló-Costa, H. L. Epistemic semantics for nonnested conditionals and nonmonotonic notions of consequence.

[17] Arló-Costa, H. L. The logic of conditionals. In Zalta [289].

[18] Avron, A. Natural 3-valued logics – Characterization and proof theory. *Journal of Symbolic Logic 56*, 1 (1991), 276–294.

[19] Ballarin, R. Modern origins of modal logic. In Zalta [289].

[20] Barwise, J. *The Situation in Logic.* CSLI Publications, Stanford, 1989.

[21] Batens, d. Paraconsistent extensional propositional logics. *Logique et Analyse 90–91* (1980), 195–234.

[22] Batens, D. Dialectical dynamics within formal logics. *Logique et Analyse 114* (1986), 161–173.

[23] Batens, D. A strengthening of the Rescher-Manor consequence relations. *Logique et Analyse 46* (2003), 289–313.

[24] Beall, J., Brady, R., Dunn, M., Hazen, A., Mares, E., Slaney, J., Meyer, R. K., Priest, G., Restall, G., Ripley, D., und Sylvan, R. On the ternary relation and conditionality. *Journal of Philosophical Logic 41* (2012), 595–612.

[25] Bennett, J. *A Philosophical Guide to Conditionals.* Clarendon Press, Oxford, 2003.

[26] Benthem, J. v. Review: B. J. Copeland, On when a semantics is not a semantics: Some reasons for disliking the Routley-Meyer semantics for Relevance Logic. *Journal of Symbolic Logic 49* (1984), 994–995.

[27] Benthem, J. v. *The Logic of Time (2nd ed.).* Kluwer, Dordrecht, 1991.

[28] Benthem, J. v. Beyond accessibility: Functional models for modal logic. In *Diamonds and Defaults*, M. de Rijke, Hg., Kluwer, Dordrecht, 1993.

[29] Benthem, J. v. Logical dynamics meets logical pluralism? *The Australasian Journal of Logic 6* (2008), 182–209.

[30] Benthem, J. v. *Modal Logic for Open Minds.* CSLI Publications, Stanford, 2010.

[31] Blackburn, P., Rijke, M. d., und Venema, Y. *Modal logic, 3rd reprint. with corr.*, vol. 53 of *Cambridge tracts in theoretical computer science.* Cambridge University Press, Cambridge, 2008.

[32] Blackburn, P., van Benthem, J., und Wolter, F., Hgg., *Handbook of Modal Logic*, vol. 3 of *Studies in Logic and Practical Reasoning.* Elsevier, Amsterdam, 2007.

[33] Boolos, G. *The Logic of Provability.* Cambridge University Press, Cambridge, 1993.

[34] Boolos, G., Burgess, J., und Jeffrey, R. *Computability and Logic (5th ed.).* Cambridge Univ. Press, Cambridge, 2007.

[35] Borges, J. L. *Labyrinthe.* Hanser, München, 1959.

[36] Brady, R. Natural deduction systems for some quantified relevant logics. *Logique et Analyse 27* (1984), 355–377.

[37] Brady, R. Normalized natural deduction systems for some relevant logics I: The logic DW. *The Journal of Symbolic Logic 71* (2006), 35–66.

[38] Brewka, G. Preferred subtheories: An extended logical framework for default reasoning. *Proceedings of the 11th International Joint Conference on Artificial Intelligence, Vol. 2* (1989), 1043–1048.

[39] Brewka, G. Cumulative default logic: In defense of nonmonotonic inference rules. *Artificial Intelligence 50*, 2 (1991), 183–205.

[40] Brewka, G. Reasoning about priorities in Default Logic. *Proceedings of the Twelfth National Conference on Artificial Intelligence (Vol. 2)* (1994), 940–945.

[41] Brewka, G., und Eiter, T. Prioritizing default logic. In *Intellectics and Computational Logic – Papers in Honour of Wolfgang Bibel*, S. Hölldobler, Hg., vol. 19 of *Applied Logic Series*. Kluwer, 2000, pp. 27–45.

[42] Brogaard, B., und Salerno, J. Remarks on counterpossibles. *Synthese* (2013), 639–660.

[43] Burgess, J. P. *Philosophical Logic*. Princeton University Press, Princeton–Oxford, 2009.

[44] Bülow, C. v. *Beweisbarkeitslogik – Gödel, Rosser, Solovay*. Logos, Berlin, 2006.

[45] Carnap, R. Modalities and quantification. *Journal of Symbolic Logic 11* (1946), 33–64.

[46] Carnap, R. *Meaning and Necessity*. Chicago University Press, Chicago und London, 1947, enlarged ed. 1956.

[47] Carnielli, W., und Pizzi, C. *Modalities and Multimodalities*. Springer, Heidelberg – New York, 2008.

[48] Carnielli, W. A., Coniglio, M. E., und Marcos, J. Logics of formal inconsistency. In Gabbay and Guenthner [93], pp. 15–107.

[49] Carnielli, W. A., und Marcos, J. A taxonomy of C-systems. In *Paraconsistency–The Logical Way to the Inconsistent,*, W. A. Carnielli, M. E. Coniglio, und I. M. L. Dottaviano, Hgg., Marcel Dekker, New York, 2002, pp. 1–94.

[50] Chagrov, A., und Zakharyaschev, M. *Modal Logic*, vol. 35 of *Oxford Logic Guides*. Oxford University Press, Oxford, 1997.

[51] Chagrova, L. An undecidable problem in correspondence theory. *Journal of Symbolic Logic 56* (1991), 1261–1272.

[52] Chellas, B. F. Basic conditional logic. *Journal of Philosophical Logic 4* (1975), 133–153.

[53] Chellas, B. F. *Modal Logic: An Introduction*. Cambridge University Press, Cambridge, 1980.

[54] Copeland, B. J. On when a semantics is not a semantics: Some reasons for disliking the Routley-Meyer semantics for Relevance Logic. *Journal of Philosophical Logic 8*, 1 (1979), 399–413.

[55] Curry, H. B. The inconsistency of certain formal logics. *Journal of Symbolic Logic 7* (1942), 115–117.

[56] da Costa, N. On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic 15* (1974), 497–510.

[57] da Costa, N., und Doria, F. On Jaśkowski's discussive logics. *Studia Logica 54* (1995), 33–66.

[58] da Costa, N. C. A., Krause, D., und Bueno, O. Paraconsistent logics and paraconsistency. In *Philosophy of Logic*, D. Jacquette, Hg., vol. 5 of *Handbook of the Philosophy of Science*. Elsevier, Amsterdam, 2006.

[59] da Costa, N. C. A., und Marconi, D. An overview of paraconsistent logic in the 80s. *The Journal of Non-Classical Logic 6*, 1 (1989), 5–32.

[60] Davis, W. Implicature. In Zalta [289].

[61] Delgrande, J., Schaub, T., und Jackson, W. Alternative Approaches to Default Logic. *Artificial Intelligence 70* (1994), 167–237.

[62] Delgrande, J., Schaub, T., Tompits, H., und Wang, K. A classification and survey of preference handling approaches in nonmonotonic reasoning. *Computational Intelligence 20* (2004), 308–334.

[63] Delgrande, J. P., und Schaub, T. Expressing default logic variants in default logic. *Journal of Logic and Computation 15* (2005), 593–621.

[64] Duca, S., und Leitgeb, H. How serious is the paradox of serious possibility? *Mind*, 481 (2012), 1–36.

[65] Dummett, M. The philosophical basis of intuitionistic logic. In *Truth and Other Enigmas*. Duckworth, London, 1973, pp. 215–247.

[66] Dummett, M. *Elements of Intuitionism*. Oxford University Press, Oxford, 1977.

[67] Dunn, J. M. Star and perp: Two treatments of negation. *Philosophical Perspectives 7* (1993), 331–357.

[68] Dunn, J. M., und Restall, G. Relevance logic. In Gabbay and Guenthner [93], pp. 117–224.

[69] Edgington, D. On conditionals. *Mind 104* (1995), 235–329.

[70] Edgington, D. Conditionals. In Goble [111], pp. 385–414.

[71] Edgington, D. Indicative conditionals. In Zalta [289].

[72] Ellis, B. Two theories of indicative conditionals. *Australasian Journal of Philosophy 62* (1984), 50–66.

[73] Epstein, R. L. *The Semantic Foundations of Logic Volume 1: Propositional Logics*. Springer, Dordrecht, 1990.

[74] Fine, K. Some connections between elementary and modal logic. *Third Scandinavian Logic Symposium* (1973), 15–31.

[75] Fine, K. An incomplete logic containing S4. *Theoria 40* (1974), 23–29.

[76] Fitting, M. Basic Modal Logic. In Gabbay et al. [94], pp. 368–448.

[77] Fitting, M. Modal proof theory. In Blackburn et al. [32], pp. 83–138.

[78] Fraassen, B. C. V. *Formal Semantics and Logic*. Macmillan, London, 1971.

[79] Frege, G. *Die Grundlagen der Arithmetik : Eine logisch mathematische Untersuchung über den Begriff der Zahl (1884)*, mit erg. Texten krit. hrsg. von Chr. Thiel. Meiner, Hamburg, 1986.

[80] Friedman, H., und Meyer, R. K. Whither relevant arithmetic? *Journal of Symbolic Logic 57*, 3 (1992), 824831.

[81] Friedrichsdorf, U. *Einführung in die klassische und intensionale Logik*. Vieweg und Teubner, Braunschweig, 1992.

[82] Fuhrmann, A. Reflective modalities and theory change. *Synthese 81* (1989), 115–134.

[83] Fuhrmann, A. Models for relevant modal logics. *Studia Logica 49* (1991), 301–315.

[84] Fuhrmann, A. Theory contraction through base contraction. *Journal of Philosophical Logic 20* (1991), 256–281.

[85] Fuhrmann, A. *An Essay on Contraction*. Studies in Logic, Language and Information. CSLI Publications, Stanford, 1997.

[86] Fuhrmann, A. Solid belief. *Theoria 63* (1997), 90–104. Auch in: LOGICA '96, hg. v. T. Childers, P. Kolář und V. Svoboda, Prag (Filosofia), 1997, 107–119.

[87] Fuhrmann, A. Theories of belief change. In *The Routledge Companion to Epistemology*, S. Bernecker und D. Pritchard, Hgg., Routledge, London, 2010.

[88] Fuhrmann, A., und Hansson, S. O. A survey of multiple contractions. *Journal of Logic, Language and Information 3* (1994), 39–76.

[89] Fuhrmann, A., und Levi, I. Undercutting and the Ramsey test for conditionals. *Synthese 101* (1994), 157–169. reprinted in *The Logic of Strategy*, C. Bicchieri, R. Jeffrey und B. Skyrms, Hgg., Oxford (University Press), 2000.

[90] Fuhrmann, A., und Mares, E. D. On S. *Studia Logica 53* (1994), 75–91.

[91] Fuhrmann, A., und Mares, E. D. A relevant theory of conditionals. *Journal of Philosophical Logic 24* (1995), 645–665.

[92] Fuhrmann, A., und Morreau, M., Hgg. *The Logic of Theory Change*. Lecture Notes in Artificial Intelligence 465. Springer, Heidelberg-Berlin-New York, 1991.

[93] Gabbay, D. M., und Guenthner, F., Hgg. *Handbook of Philosophical Logic (2nd Ed.)*. Springer, 2001ff.

[94] Gabbay, D. M., Hogger, C., und Robinson, J., Hgg. *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 1: Logical Foundations. Clarendon Press, Oxford, 1993.

[95] Garcia-Carpintero, M., und Macià, J., Hgg. *Two-Dimensional Semantics*. Oxford University Press, Oxford, 2006.

[96] Gärdenfors, P. Conditionals and changes of belief. *Acta Philosophica Fennica 30* (1978), 381–404.

[97] Gärdenfors, P. Rules for rational changes of belief. *Philosophical Essays dedicated to Lennart Åqvist on his fiftieth birthday* (1982), 88–101.

[98] Gärdenfors, P. Epistemic importance and minimal changes of belief. *Australasian Journal of Philosophy 62*, 2 (1984), 136–157.

[99] Gärdenfors, P. Belief revision and the Ramsey Test for conditionals. *The Philosophical Review 95*, 1 (1986), 81–93.

[100] Gärdenfors, P. *Knowledge in Flux.* MIT Press, Cambridge, Mass., 1988.

[101] Gärdenfors, P., und Makinson, D. Revisions of Knowledge Systems Using Epistemic Entrenchment. In *Proceedings of the 2Nd Conference on Theoretical Aspects of Reasoning About Knowledge*, TARK '88. Morgan Kaufmann Publishers Inc., 1988, pp. 83–95.

[102] Gärdenfors, P., und Makinson, D. Nonmonotonic inference based on expectations. *Artificial Intelligence 65* (1994), 197–245.

[103] Gärdenfors, P., und Rott, H. Belief revision. In *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 4: Epistemic and Temporal Reasoning*, D. Gabbay, C. Hogger, und J. Robinson, Hgg., Oxford University Press, Oxford, 1995, pp. 35–132.

[104] Garson, J. W. Two new interpretations of modality. *Logique et Analyse 15* (1972), 443–459.

[105] Gelfond, M., Lifschitz, V., Przymusinska, H., und Truszczynski, M. Disjunctive defaults. In *Proceedings of International Conference on Principles of Knowledge Representation and Reasoning (KR)* (1991), J. Allen, R. Fikes, und E. Sandewall, Hgg., pp. 230–237.

[106] Gentzen, G. Untersuchungen über das logische Schließen. *Mathematische Zeitschrift 39* (1934–35), 176–210, 405–431.

[107] Gerson, M. The inadequacy of the neighbourhood semantics for modal logic. *The Journal of Symbolic Logic 40* (1975), 141–148.

[108] Gibbard, A. Two recent theories of conditionals. In Harper et al. [130].

[109] Gillies, A. S. What might be the case after a change in view. *Journal of Philosophical Logic 35* (2006), 117–145.

[110] Ginsberg, M. L., Hg. *Readings in Nonmonotonic Reasoning.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.

[111] Goble, L., Hg. *The Blackwell Guide to Philosophical Logic.* Blackwell, Oxford, 2001.

[112] Gödel, K. Eine Interpretation des intuitionistischen Aussagenkalküls. *Ergebnisse eines mathematischen Kolloquiums, Heft 4* (1933), 39–40.

[113] Goldblatt, R. *Logics of Time and Computation.* CSLI Publications, Stanford, 1987.

[114] Goldblatt, R. *Mathematics of Modality.* CSLI Publications, Stanford, 1993.

[115] Goldblatt, R. Equivalent beliefs in Dynamic Doxastic Logic. In *Krister Segerberg on Logic of Actions*, R. Trypuz, Hg., Springer, Dordrecht, 2014, pp. 179–207.

[116] Goldblatt, R. I. Semantic Analysis of Orthologic. *Journal of Philosophical Logic 3*, 1/2 (1974), 19–35.

[117] Goranko, V., und Rumberg, A. Temporal Logic. *Stanford Encyclopedia of Philosophy* (2022).

[118] Grice, P. Logic and conversation. In *Syntax and semantics. Vol. 3: Speech Acts*, P. Cole und J. Morgan, Hgg., Academic Press, New York, 1975, pp. 41–58.

[119] Grice, P. *Studies in the Ways of Words.* Harvard University Press, Harvard, 1989.

[120] Grove, A. Two modellings for theory change. *Journal of Philosophical Logic 17* (1988), 157–170.

[121] Hájek, A. Probabilities of conditionals revisited. *Journal of Philosophical Logic 18* (1989), 423–428.

[122] Hájek, A. Probability, logic, and probability logic. In Goble [111].

[123] Hájek, A. The fall of "Adams' Thesis"? *Journal of Logic, Language and Information 21*, 2 (2012), 145–161.

[124] Hakli, R., und Negri, S. Does the Deduction Theorem fail for modal logic? *Synthese 187*, 3 (2012), 849–867.

[125] Hansson, S. O. New operators for theory change. *Theoria 55* (1989), 114–132.

[126] Hansson, S. O. In defense of base contraction. *Synthese 91* (1992), 239–245.

[127] Hansson, S. O. Kernel contraction. *The Journal of Symbolic Logic 59* (1994), 845–859.

[128] Hansson, S. O. A survey of non-prioritized belief revision. *Erkenntnis 50* (1999), 413–427.

[129] Hansson, S. O. *A Textbook of Belief Dynamics.* Kluwer, Dordrecht, 1999.

[130] Harper, W., Stalnaker, R., und Pearce, G., Hgg. *Ifs: Conditionals, Belief, Decision, Chance, and Time.* Reidel, Dordrecht, 1981.

[131] Harper, W. L. Rational conceptual change. *PSA 2* (1976), 462–494.

[132] Hazen, A., Rin, B., und Wehmeier, K. Actuality in propositional modal logic. *Studia Logica 101* (2013), 487503.

[133] Heyting, A. *Intuitionism. An Introduction.* North-Holland, Amsterdam, 1956.

[134] Hilbert, D., und Bernays, P. *Grundlagen der Mathematik.* Springer, Berlin, 1970.

[135] Hilpinen, R. *Philosophical Logic.* Reidel Publishing Company, Dordrecht, 1969.

[136] Hilpinen, R., Hg. *Deontic Logic: Introductory and Systematic Readings.* Reidel, Dordrecht, 1981.

[137] Hintikka, J. *Knowledge and Belief.* Cornell University Press, Ithaca, NY, 1962.

[138] Horty, J. Defaults with priorities. *Journal of Philosophical Logic 36*, 4 (2007), 367–413.

[139] Horty, J. Reasons as defaults. *Philosophers' Imprint 7*, 3 (2007), 1–28.

[140] Horty, J. F. *Reasons as Defaults.* Oxford University Press, New York, 2012.

[141] Hughes, G., und Cresswell, M. *A New Introduction to Modal Logic.* Routledge, London, 1996.

[142] Humberstone, L. Two-dimensional adventures. *Philosophical Studies 118*, 1 (Mar. 2004), 17–65.

[143] Humberstone, L. *The Connectives.* MIT Press, Cambridge Mass., 2011.

[144] Humberstone, L. *Philosophical Applications of Modal Logic.* College Publications, London, 2015.

[145] Hyde, D. From heaps and gaps to heaps of gluts. *Mind 106* (1997), 440–460.

[146] Jackson, F. On assertion and indicative conditionals. *Philosophical Review 88* (1979), 565–589; also in [150].

[147] Jackson, F. Conditionals and possibilia. *Proceedings of the Aristotelian Society 81* (1980), pp. 125–137.

[148] Jackson, F. *Conditionals.* Blackwell, Oxford, 1987.

[149] Jackson, F., Hg. *Conditionals.* Oxford University Press, Oxford, 1991.

[150] Jackson, F. *Mind, Matter and Conditionals.* Routledge, London, 1998.

[151] Jackson, F. Postscript on truth conditions and assertability. In *Mind, Matter and Conditionals* [150], pp. 51 – 54.

[152] Jacquette, D., Hg. *A Companion to Philosophical Logic.* Blackwell, Oxford, 2002.

[153] Janhunen, T. Classifying semi-normal default logic on the basis of its expressive power. In *Proceedings of the Fifth International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR99)*, vol. 1730 of *Lecture Notes in Artificial Intelligence.* Springer, 1999.

[154] Jankov, V. Some superconstructive propositional calculi. *Soviet Mathematics (Doklady) 4* (1963), 1103–1105.

[155] Jaśkowski, S. Rachunek zdań dla systemów dedukcyjnych sprzecznych. *Studia Societatis Scientiarun Torunesis (Sectio A) 1* (1948), 55–77.

[156] Jaśkowski, S. Propositional calculus for contradictory deductive systems (Engl. tansl. of [155]). *Studia Logica 24* (1969), 143–157.

[157] Johansson, I. Der Minimalkalkül, ein reduzierter intuitionistischer Formalismus. *Compositio Mathematica 4* (1937), 119–136.

[158] Jónsson, B., und Tarski, A. Boolean algebras with operators, I. *American Journal of Mathematics 73* (1952), 891–939.

[159] Kanger, S. *Provability in Logic.* Almqvist & Wiksell, Stockholm, 1957.

[160] Kaplan, D. On the logic of demonstratives. *Journal of Philosophical Logic 8* (1979), 81–98.

[161] Kleene, S. *Introduction to Metamathematics.* North-Holland, Amsterdam-New York, 1952.

[162] Kracht, M. *Tools and Techniques in Modal Logic*, vol. 142 of *Studies in Logic and the Foundations of Mathematics*. Elsevier, Amsterdam, 1999.

[163] Kraus, S., Lehmann, D., and Magidor, M. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence 44* (1990), 167–207.

[164] Kripke, S. A completeness theorem in modal logic. *Journal of Symbolic Logic 24* (1959), 1–14.

[165] Kripke, S. A. Semantical analysis of modal logic, I. Normal modal propositional calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik 9* (1963), 67–96.

[166] Kripke, S. A. Semantical analysis of intuitionistic logic. In *Formal Systems and Recursive Functions* (Amsterdam, 1965), J. Crossley und M. Dummett, Hgg., North-Holland, pp. 92–130.

[167] Kripke, S. A. On two paradoxes of knowledge. In *Philosophical Troubles: Collected Papers, Volume 1.* Oxford University Press, Oxford, 2011.

[168] Kyburg, H. E. Probability, rationality, and the rule of detachment. In *Proceedings of the 1964 International Congress for Logic, Methodology, and Philosophy of Science*, Y. Bar-Hillel, Hg., North-Holland, Amsterdam, 1965, pp. 301–310.

[169] Lasswitz, K. Die Universalbibliothek. *Ostdeutsche Allgemeine Zeitung 18. Dez. 1904* (1904).

[170] Lemmon, E. Algebraic semantics for modal logics, I–II. *Journal of Symbolic Logic* (1966), 46–65, 191–218.

[171] Lemmon, E. *An Introduction to Modal Logic (The "Lemmon Notes")*, ed. by K. Segerberg, vol. 11 of *American Philosophical Quarterly Monograph Series.* Blackwell, Oxford, 1977.

[172] Levi, I. *The Enterprise of Knowledge.* MIT Press, Cambridge, Mass., 1980.

[173] Levi, I. Truth, fallibility and the growth of knowledge. In *Language, Logic and Method*, R. S. Cohen und M. W. Wartofsky, Hgg., Springer, Dordrecht, 1983, pp. 153–174.

[174] Levi, I. *The Fixation of Belief and Its Undoing.* Cambridge Univ. Press, Cambridge, 1991.

[175] Levi, I. *For the Sake of the Argument.* Cambridge Univ. Press, Cambridge, 1996.

[176] Levi, I. *Mild Contractions.* Cambridge Univ. Press, Cambridge, 2004.

[177] Lewis, C., und Langford, C. *Symbolic Logic.* Dover, New York, 1932.

[178] Lewis, C. I. Implication and the algebra of logic. *Mind 21* (1912), 522–531.

[179] Lewis, C. I. The calculus of strict implication. *Mind 23* (1914), 240–247.

[180] Lewis, D. How to define theoretical terms. *Journal of Philosophy 67* (1970), 427–446.

[181] Lewis, D. *On the Plurality of Worlds.* Blackwell, Oxford, 1986.

[182] Lewis, D. *Counterfactuals.* Blackwell, Oxford, 1973.

[183] Lewis, D. Probabilities of conditionals and conditional probabilities. *The Philosophical Review 85* (1976), 297–315.

[184] Lewis, D. Attitudes de dicto and de se. *The Philosophical Review 88* (1979).

[185] Lewis, D. Logic for equivocators. *Nous 16* (1982), 431–441.

[186] Lewis, D. Postscript to "Probabilities of conditionals and conditional probabilities". In *Philosophical Papers II.* Oxford University Press, Oxford, 1986, pp. 152–156.

[187] Lindström, S. The Ramsey Test and the indexicality of conditionals: A proposed resolution of Gärdenfors' paradox. In Fuhrmann und Morreau [92], pp. 208–228.

[188] Lindström, S., und Segerberg, K. Modal logic and philosophy. In *Handbook of Modal Logic*, P. Blackburn und J. van Benthem, Hgg., Elsevier, 2007.

[189] Łukaszewicz, W. Considerations on default logic: An alternative approach. *Computational Intelligence 4* (1988), 1–16.

[190] Löb, M. Solution to a problem of Leon Henkin. *Journal of Symbolic Logic 20* (1955), 115–118.

[191] Makinson, D. How meaningful are modal operators? *Australasian Journal of Philosophy 44* (1966), 331–337.

[192] Makinson, D. On some completeness theorems in modal logic. *Zeitschr. f. math. Logik und Grundlagen d. Math. 12* (1966), 379–384.

[193] Makinson, D. A generalisation of the concept of a relational model for modal logic. *The Journal of Symbolic Logic 38* (1973), 330–335.

[194] Makinson, D. How to give it up: A survey of some formal aspects of the logic of theory change. *Synthese 62* (1985), 347–363.

[195] Makinson, D. On the status of the postulate of recovery in the logic of theory change. *Journal of Philosophical Logic 16* (1987), 383–394.

[196] Makinson, D. Five faces of minimality. *Studia Logica 52* (1993), 339–379.

[197] Makinson, D. General patterns in nonmonotonic reasoning. In *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 4: Epistemic and Temporal Reasoning*, D. Gabbay, C. Hogger, und J. Robinson, Hgg., Oxford University Press, Oxford, 1994, pp. 35–110.

[198] Makinson, D. Screened revision. *Theoria 63* (1997), 14–23.

[199] Makinson, D. Ways of doing logic: What was different about AGM 1985? *Journal of Logic and Computation 13* (2003), 3–13.

[200] Makinson, D. *Bridges from Classical to Nonmonotonic Logic*. King's College Publications, London, 2005.

[201] Makinson, D. How to go nonmonotonic. In *Handbook of Philosophical Logic*, D. Gabbay und F. Guenthner, Hgg., vol. 12 of *Handbook of Philosophical Logic*. Springer, 2005, pp. 175–278.

[202] Makinson, D. *Sets, Logic and Maths for Computing.* Springer, London, 2008, rev. and exp. ed. 2020.

[203] Makinson, D. Relevance logic as a conservative extension of classical logic. In *David Makinson on Classical Methods for Non-Classical Problems*. Springer Netherlands, dec 2013, pp. 383–398.

[204] Makinson, D. The relevance logic program: failed or just stalled? In *Sets, Logic and Maths for Computing* [202], p. as ch. 11 in the 2020 edition.

[205] Makinson, D. Relevance-sensitive truth-trees. In *Alasdair Urquhart on Nonclassical and Algebraic Logic and Complexity of Proofs*, I. Düntsch und E. Mares, Hgg., Springer, Cham, Nov. 2022, pp. 23–65.

[206] Makinson, D., und Gärdenfors, P. Relations between the logic of theory change and nonmonotonic logic. In Fuhrmann und Morreau [92], pp. 183–205.

[207] Makinson, D., und van der Torre, L. Input/output logics. *Journal of Philosophical Logic 29* (2000), 383–408.

[208] Marcos, J. *Logics of Formal Inconsistency.* PhD thesis, Universidade Técnica de Lisboa, 2005.

[209] Mares, E. *Relevant Logic. A Philosophical Interpretation.* Cambridge University Press, Cambridge, 2004.

[210] Mares, E., and Meyer, R. Relevant logics. In Goble [111], pp. 280–308.

[211] McKinsey, J., und Tarski, A. Some theorem about the sentential calculi of Lewis and Heyting. *The Journal of Symbolic Logic 13* (1948), 1–15.

[212] Meyer, R. K. New axiomatics for relevant logics, I. *Journal of Philosophical Logic 3* (1974), 53–86.

[213] Meyer, R. K., und Routley, R. Classical relevant logics. I. *Studia Logica 32* (1973), 51–68.

[214] Meyer, R. K., und Routley, R. Classical relevant logics. II. *Studia Logica 33* (1974), 183–194.

[215] Moore, R. Semantical considerations on nonmonotonic logic. *Artificial Intelligenc 25* (1985), 75–94.

[216] Nasieniewski, M., und Pietruszczak, A. A method of menerating modal logics defining D_2. *Studia Logica 97* (2011), 161–182.

[217] Omori, H., und Alama, J. Axiomatizing Jaśkowski's discussive logic D_2. *Studia Logica 106* (2018), 1168–1180.

[218] Pacuit, E. *Neighborhood Semantics for Modal Logic.* Springer, Heidelberg - New York, 2017.

[219] Poole, D. A logical framework for default reasoning. *Artificial Intelligence 36* (1988), 27–47.

[220] Poole, D. Default logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 4: Epistemic and Temporal Reasoning*, D. Gabbay, C. Hogger, und J. Robinson, Hgg., Oxford University Press, Oxford, 1994, pp. 189–215.

[221] Priest, G. Logic of paradox. *Journal of Philosophical Logic 8* (1979), 219–241.

[222] Priest, G. *Beyond the Limits of Thought.* 2nd ed., Cambridge University Press, Cambridge, 2002.

[223] Priest, G. Paraconsistent logic. In Gabbay and Guenthner [93], pp. 287–393.

[224] Priest, G. *Towards Non-Being: the Semantics and Metaphysics of Intentionality.* Oxford University Press, Oxford, 2005.

[225] Priest, G. *In Contradiction: A Study of the Transconsistent.* 2nd ed., Oxford University Press, Oxford, 2006.

[226] Priest, G. Paraconsistency and dialetheism. In *Handbook of the History of Logic*, D. Gabbay und J. Woods, Hgg., vol. 8. North Holland, Amsterdam, 2007, pp. 129–204.

[227] Priest, G. *An Introduction to Non-Classical Logic: From If to Is.* 2nd ed., Cambridge University Press, Cambridge, 2008.

[228] Priest, G. *Einführung in die nicht-klassische Logik.* Mentis, Paderborn, 2008.

[229] Priest, G., und Routley, R. Systems of paraconsistent logic. In Priest et al. [230], pp. 151–186.

[230] Priest, G., Routley, R., und Norman, J., Hgg. *Paraconsistent Logic: Essays on the Inconsistent.* Philosophia Verlag, München, 1989.

[231] Prior, A. *Time and Modality.* Oxford University Press, Oxford, 1957.

[232] Prior, A. N. *Past, Present and Future.* Clarendon Press, Oxford, 1967.

[233] Quine, W. V. O. The problem of interpreting modal logic. *The Jounal of Symbolic Logic 12* (1947), 43–48.

[234] Quine, W. V. O. *Methods of Logic.* Harvard University Press, Cambridge, Mass., 1950.

[235] Quine, W. V. O. *The Ways of Paradox.* Random House, New York, 1966.

[236] Quine, W. V. O. On the reasons for indeterminacy of translation. *The Journal of Philosophy 67*, 6 (1970), 178–183.

[237] QUINE, W. V. O. *Philosophy of Logic.* Prentice Hall, Englewood Cliffs, N.J., 1970.

[238] QUINE, W. V. O., UND CARNAP, R. *Dear Carnap, Dear Van : The Quine–Carnap Correspondence and Related Work*, ed. by R. Creath, Univ. of California Press, Berkeley, 1990.

[239] RAMSEY, F. P. *The Foundations of Mathematics.* Routledge and Kegan Paul, Oxford, 1931.

[240] READ, S. *Relevant Logic: A Philosophical Examination of Inference.* Blackwell, Oxford, 1988.

[241] READ, S. L. *Thinking About Logic.* Oxford University Press, Oxford, 1994.

[242] REIDHAAR-OLSON, L. A new proof of the fixed-point theorem of provability logic. *Notre Dame J. Formal Logic 31* (1990), 37–43.

[243] REITER, R. A logic for default reasoning. *Artificial Intelligence 13* (1980), 81–132; reprinted in [110, 68–93].

[244] RESCHER, N., UND BRANDOM, R. *The Logic of Inconsistency.* Blackwell, Oxford, 1980.

[245] RESCHER, N., UND MANOR, R. On inference from inconsistent premises. *Theory and Decision* (1970), 179–217.

[246] RESTALL, G. Negation in Relevant Logics (how I stopped worrying and learned to love the Routley star). In *What is Negation?*, D. Gabbay und H. Wansing, Hgg., vol. 13 of *Applied Logic Series.* Kluwer Academic Publishers, Amsterdam, 1999, pp. 53–76.

[247] RESTALL, G. *An Introduction to Substructural Logics.* Routledge, 2000.

[248] RESTALL, G. Defining double negation elimination. *Logic Journal of IGPL 8* (2000), 853–860.

[249] ROTT, H. Conditionals and theory change: Revisions, expansions and additions. *Synthese 81* (1989), 91–113.

[250] ROTT, H. On the logic of theory change: More maps between different kinds of contraction functions. In *Belief Revision*, P. Gärdenfors, Hg., Cambridge Univ. Press, Cambridge, 1992, pp. 122–141.

[251] ROTT, H. *Change, Choice and Inference.* Clarendon Press, Oxford, 2001.

[252] Rott, H., und Hansson, S. O. Safe contraction revisited. In *David Makinson on Classical Methods for Non-Classical Problems*, S. O. Hansson, Hg., Springer, Dordrecht, 2014, pp. 35–70.

[253] Rott, H., und Pagnucco, M. Severe withdrawal (and recovery). *Journal of Philosophical Logic 28*, 5 (1999), 501–547.

[254] Routley, R. *Relevant Logics and Their Rivals. Part I: The Basic Philosophical and Semantical Theory.* Ridgeview Publishing, Atascadero, 1982.

[255] Sangiorgi, D. On the origins of bisimulation and coinduction. *ACM Trans. Program. Lang. Syst. 31*, 4 (2009), 15:1–15:41.

[256] Schaub, T. On constrained default theories. In *European Conference on Artificial Intelligence (ECAI)* (1992), B. Neumann, Hg., Wiley, pp. 304–308.

[257] Schotch, P., Brown, B., und Jennings, R., Hgg. *On Preserving: Essays on Preservationism and Paraconsistent Logic.* University of Toronto Press, Toronto, 2009.

[258] Schotch, P., und Jennings, R. Inference and necessity. *Journal of Philosophical Logic 9* (1980), 327–340.

[259] Schurz, G. Carnaps modal logic. In *Zwischen traditioneller und moderner Logik*, W. Stelzner and M. Stockler, Hgg., Mentis, Paderborn, 2001, pp. 365–380.

[260] Segerberg, K. *An Essay in Classical Modal Logic.* Filosofiska Institut, Uppsala Universitet, Uppsala, 1971.

[261] Segerberg, K. Two-dimensional modal logic. *Journal of Philosophical Logic 2* (1973), 77–69.

[262] Segerberg, K. A note on an impossibility theorem of Gärdenfors. *Noûs 23*, 3 (1989), 351–354.

[263] Segerberg, K. Notes on conditional logic. *Studia Logica 48* (1989), 157–168.

[264] Shehtman, V. On strong neighbourhood completeness of propositional logics, part I. In *Advances in Modal Logic*, M. Kracht, M. de Rijke, H. Wansing, und M. Zakharyashev, Hgg., CSLI Publications, Stanford, 1997.

[265] Shehtman, V. On neighbourhood semantics thirty years later. In *We Will Show Them! Essays in Honour of Dov Gabbay*, S. Artemov et al., Hgg., College Publications, London, 2005, pp. 663–692.

[266] Shoham, Y. *Reasoning About Change.* MIT Press, Cambridge Mass., 1988.

[267] Slaney, J. A general logic. *Australasian Journal of Philosophy 68* (1990), 74–88.

[268] Slaney, J. K. Reduced models for relevant logics without WI. *Notre Dame Journal of Formal Logic 28* (1987), 395–407.

[269] Smith, P. *An Introduction to Gödel's Theorems.* Cambridge Univ. Press, Cambridge, 2013.

[270] Smith, P. *Gödel Without (Too Many) Tears (2nd ed.).* Logic Matters, 2022.

[271] Smoryński, C. *Self-Reference and Modal Logic.* Springer, New York, 1985.

[272] Stalnaker, R. Probability and conditionals. *Philosophy of Science 37* (1970), 64–80.

[273] Stalnaker, R. A defense of Conditional Excluded Middle. In Harper et al. [130], pp. 87–104.

[274] Stalnaker, R. Nonmonotonic consequence relations. *Fundamenta Informaticae 21* (1994), 7–21.

[275] Stalnaker, R. C. A theory of conditionals. In Harper et al. [130], pp. 41–55.

[276] Sylvan, R. *Sociative Logics and Their Applications.* Ashgate, Aldershot, 2000.

[277] Sylvan (Routley), R. *Relevant Logics and Their Rivals. Part II.* Ashgate, Aldershot, 2003.

[278] Tennant, N. Perfect validity, entailment and paraconsistency. *Studia Logica 43* (1984), 179–198.

[279] Thomason, R. H. Combinations of tense and modality. In Gabbay and Guenthner [93], pp. vol. 7, 205–234.

[280] Thomason, S. An incompleteness theorem in modal logic. *Theoria 40* (1974), 150–158.

[281] Thomason, S. K. New representation of S5. *Notre Dame Journal of Formal Logic 14* (1973), 281–284.

[282] Unger, P. There are no ordinary things. *Synthese 41* (1979), 117–154.

[283] Urbas, I. Paraconsistency and the C-systems of da Costa. *Notre Dame J. Formal Logic 30* (1989), 583–597.

[284] Urchs, M. Discursive logic: Towards a logic of rational discourse. *Studia Logica 54*, 2 (1995), 231–249.

[285] Walton, D. N. Philosophical basis of relatedness logic. *Philosophical Studies 36*, 2 (1979), 115–136.

[286] Wheeler, S. Reference and Vagueness. *Synthese 30* (1975), 367–379.

[287] Wright, G. H. v. Deontic Logic. *Mind 60* (1951), 1–15.

[288] Wölfl, S. *Kombinierte Zeit- und Modallogik.* Logos, Berlin, 1999.

[289] Zalta, E. N., Hg. *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, Stanford.

[290] Zanardo, A. Axiomatization of 'Peircean' branching-time logic. *Studia Logica 49* (1990), 183–195.

[291] Zanardo, A. A note about the axioms for branching-time logic. *Notre Dame Journal of Formal Logic 33* (1992), 225–228.

[292] Zanardo, A. Branching-time logic with quantification over branches: The point of view of modal logic. *Journal of Symbolic Logic 61* (1996), 1–39.

[293] Zanardo, A. Modalities in temporal logic. *Humana Mente 3* (2009), 1–15.

[294] Zinke, A. *The Metaphysics of Logical Consequence.* Klostermann, Frankfurt, 2018.

INDEX

„Philosophische Logik“ bezeichnet einen Fundus von Theorien und Methoden, ohne die weite und zentrale Teile der heutigen Philosophe – in beinahe wörtlichem Sinne – gar nicht denkbar sind. Im Wechselspiel mit ihren Anwendungen hat sich dabei der Begriff der Logik selbst weiterentwickelt. An den Schnittstellen zu Nachbardisziplinen der Philosophie, wie der Informatik und Linguistik, hat sich die Entwicklung besonders rasant vollzogen und ist wiederum fruchtbar für die philosophische Theorienbildung geworden. Fuhrmanns Frankfurter Vorlesungen versuchen in dieses moderne Verständnis philosophischer Logik einzuführen. Das Buch ist als Lehrbuch geeignet, leitet den Leser jedoch auch zur kritischen Beurteilung formaler Theorien in der Philosophie an. Zentrale Themen werden kapitelübergreifend entwickelt. Dennoch sind die Kapitel in sich abgeschlossen und können so als gründliche Einführungen in ausgewählte Theorien gelesen werden.

Grundlagen — Zeitlogik — Modallogik
Konditionale — Parakonsistente Logik — Relevanzlogik
Anfechtbares Schließen — Überzeugungswandel

André Fuhrmann ist Professor für Theoretische Philosophie mit Schwerpunkt Logik an der Goethe-Universität Frankfurt am Main.

Logos Verlag Berlin

ISBN 978-3-8325-5654-9

https://www.logos-verlag.de/oekobuch