Franz von Kutschera

Normative Präferenzen und bedingte Gebote

Seit dem Erscheinen der Arbeiten von S. Halldén (57)* und von G. H. von Wright (63) ist die Logik der Präferenzbegriffe in der Literatur viel diskutiert worden[1]. Ein Präferenzbegriff ist eine Relation $A \leq \cdot B$, die besagt, daß die durch den Satz A ausgedrückte Proposition höchstens so hoch bewertet wird wie die durch den Satz B ausgedrückte Proposition. Es hat sich gezeigt, daß es eine Fülle von Möglichkeiten gibt, solche Relationen zu interpretieren, und daß man dabei zu sehr verschiedenen Eigenschaften der Relation $A \leq \cdot B$ gelangt. An die Diskussion von Präferenzen haben sich dann zahlreiche Untersuchungen angeschlossen mit dem Ziel, bedingte Obligationen der Form $O(A/B)$ – unter der Bedingung, daß B, ist es geboten, daß A – auf der Basis normativer Präferenzrelationen zu definieren. Auch hier haben verschiedene Ansätze zu sehr unterschiedlichen Systemen der Logik bedingter Obligationen geführt.

Im folgenden werden die wichtigsten Ansätze zur Bestimmung von Präferenzrelationen und bedingten Obligationen diskutiert. Der Kürze wegen beschränken wir uns, wie die meisten Arbeiten zu unserem Thema, auf die Sprache der Aussagenlogik, sowie auf die Untersuchung nichtiterierter Anwendungen deontischer Operatoren[2]. Die Erörterungen führen 1. dazu, daß ein probabilistischer Präferenzbegriff intuitiv am angemessensten erscheint, für den es jedoch im Rahmen einer aussagenlogischen Sprache bisher keine vollständige Formulierung gibt. Sie ergeben 2. eine Formalisierung eines intuitiv akzeptablen Begriffs bedingter Obligationen, der jedoch nicht direkt mit Präferenzen für Propositionen zusammenhängt. Als formal befriedigender, intuitiv aber nicht unproblemati-

* Die Zahlen weisen auf das Literaturverzeichnis am Ende dieses Kapitels, Seite 165 hin.

[1] Für historische Bemerkungen zur Präferenzlogik und eine Bibliographie dazu vgl. Rescher (67).

[2] Vgl. dazu Kutschera (73) (im folgenden zitiert als »NWE«), Abschnitt 1.5.

scher Ersatz für eine umfassende Formalisierung von Präferenzen und bedingten Geboten wird endlich 3. eine Präferenzlogik angegeben, in der sich bedingte Obligationen explizit definieren lassen.

1 Die Sprache Δ

Die Sprache der im folgenden untersuchten Systeme enthält abzählbar unendlich viele Satzkonstanten. Die *Sätze* von Δ werden so bestimmt:

1.1 a) Jede Satzkonstante von Δ ist ein Satz von Δ.

b) Ist A ein Satz von Δ, so auch $\neg A$.

c) Sind A und B Sätze von Δ, so ist auch $(A \supset B)$ ein Satz von Δ.

d) Ist A ein Satz von Δ, in dem deontische Operatoren (O und $\leq\cdot$) nicht vorkommen, so ist auch $O(A)$ ein Satz von Δ.

e) Sind A und B Sätze von Δ, in denen deontische Operatoren nicht vorkommen, so ist auch $O(A/B)$ ein Satz von Δ.

f) Sind A und B Sätze von Δ, in denen deontische Operatoren nicht vorkommen, so ist auch $(A \leq\cdot B)$ ein Satz von Δ.

Δ_0 sei die aussagenlogische Teilsprache von Δ, deren Sätze allein mit 1.1-a, b, c gebildet werden. Δ_1 sei die Teilsprache von Δ, deren Sätze allein mit 1.1-a, b, c, f gebildet werden. Δ_2 sei die Teilsprache von Δ, deren Sätze allein nach 1.1-a, b, c, d gebildet werden. Und Δ_3 sei die Teilsprache von Δ, deren Sätze allein mit 1.1-a, b, c, d, e gebildet werden.

1.2 Definitionen:

a) $A \vee B := \neg A \supset B$

b) $A \wedge B := \neg(\neg A \vee \neg B)$

c) $A \equiv B := (A \supset B) \wedge (B \supset A)$

d) $A <\cdot B := (A \leq\cdot B) \wedge \neg(B \leq\cdot A)$

e) $A =\cdot B := (A \leq\cdot B) \wedge (B \leq\cdot A)$

f) $V(A) := O(\neg A)$ und $V(A/B) := O(\neg A/B)$

g) $E(A) := \neg O(\neg A)$ und $E(A/B) := \neg O(\neg A/B)$

h) $I(A) := \neg O(A) \wedge \neg O(\neg A)$ und $I(A/B) := \neg O(A/B) \wedge \neg O(\neg A/B)$.

Wir schreiben $A_1, \ldots, A_n \rightarrow B$, falls der Schluß von $A_1, \ldots, A_n$ auf B aussagenlogisch gültig ist. Gilt $A \rightarrow \neg B$, so schreiben wir statt $A \vee B$ auch $A + B$. Diese Schreibweise soll umgekehrt auch immer voraussetzen, daß $A \rightarrow \neg B$ gilt. Klammern lassen wir weg, sofern sie zur Abgrenzung des Anwendungsbereichs der Operatoren nicht nötig sind, wobei wir festlegen, daß in der Folge der aussagenlogischen Operatoren $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$ jeder links von einem Operator stehende Operator stärker bindet als dieser.

Um die folgenden metasprachlichen Festlegungen kurz und übersichtlich formulieren zu können, verwenden wir neben den üblichen mengentheoretischen Symbolen gelegentlich auch die logischen Symbole von Δ, sowie die Existenz- und Allquantoren V und Λ als metasprachliche Zeichen. Verwechselungen objekt- und metasprachlicher Sätze können dabei nicht auftreten.

2 Präferenzen zwischen Welten

Der einfachste und naheliegendste Weg zur Festlegung von Präferenzen zwischen Propositionen ist es, von einer 2-stelligen Präferenzrelation zwischen Welten auszugehen. Es seien die $i \in I$ Indices für Welten, die bzgl. der realen Welt als mögliche Welten in Betracht gezogen werden und auf I sei eine Präferenzrelation $\cdot \leq$ definiert. Diese Präferenzrelation soll eine *partielle Quasiordnung* darstellen, d.h. sie soll reflexiv und transitiv sein. Für alle i, j, k aus I soll also gelten:

2.1 a) $i \cdot \leq i$
b) $i \cdot \leq j \wedge j \cdot \leq k \supset i \cdot \leq k$.

Eine *totale Quasiordnung* liegt vor, wenn alle $i, j \in I$ auch vergleichbar sind, d.h. wenn gilt:

c) $i \cdot \leq j \vee j \cdot \leq i$[3].

Wir definieren in Analogie zu 1.2—d, e

2.2 a) $i \cdot < j := i \cdot \leq j \wedge \neg (j \cdot \leq i)$[4]
b) $i \cdot = j := i \cdot \leq j \wedge j \cdot \leq i$.

Der zu definierenden Relation $\cdot \leq$ zwischen Propositionen entspricht nun nicht direkt die Relation $\cdot \leq$ auf I, sondern eine Relation zwischen Mengen von Welten aus I. Die Proposition,

[3] Aus (c) folgt dann (a).
[4] Die Definition $i \cdot < j := \neg (j \cdot \leq i)$ würde dazu führen, daß die Trichotomie $i = \cdot j \vee i \cdot < j \vee j \cdot < i$ gilt, d.h.daß eine totale Quasiordnung vorliegt.

die ein Satz A ausdrückt, ist ja darzustellen durch die Menge $[A]$ der Welten aus I, in denen A wahr ist. Man muß also eine Präferenzrelation für die Mengen $[A]$ zu den Sätzen A von Δ_0 aufgrund der Relation $\cdot \leq$ erklären.

Es gibt nun eine Reihe von Möglichkeiten, eine Präferenzordnung $\leq \cdot$ für Propositionen mithilfe der Relation $\cdot \leq$ zu definieren. Von der Relation $\leq \cdot$ verlangen wir erstens, daß sie eine partielle Quasiordnung darstellt. Intuitiv ist es allerdings oft durchsichtiger, anstelle der Relation $\leq \cdot$ zunächst zwei Relationen $< \cdot$ und $= \cdot$ zu definieren. Diese bilden eine partielle Quasiordnung, wenn gilt:

2.3 a) $A = \cdot A$

b) $A = \cdot B \supset B = \cdot A$

c) $A = \cdot B \wedge B = \cdot C \supset A = \cdot C$

d) $\neg (A < \cdot A)$

e) $A < \cdot B \wedge B < \cdot C \supset A < \cdot C$

f) $A = \cdot B \supset (A < \cdot C \supset B < \cdot C)$

g) $A = \cdot B \supset (C < \cdot A \supset C < \cdot B)$

Eine totale Quasiordnung liegt vor, wenn zudem gilt

h) $A = \cdot B \vee A < \cdot B \vee B < \cdot A$.

Man definiert dann:

2.4 $A \leq \cdot B := A < \cdot B \vee A = \cdot B$.

Von der Relation $\leq \cdot$ verlangen wir zweitens, daß sie das *schwache Mittelwertprinzip* erfüllt:

2.5 a) $A \leq \cdot B \supset (\neg (A \leq \cdot A + B) \wedge \neg (A + B \leq \cdot A) \vee A \leq \cdot A + B) \wedge (\neg (B \leq \cdot A + B) \wedge \neg (A + B \leq \cdot B) \vee A + B \leq \cdot B)$.[5]

Ist $A + B$ sowohl mit A wie mit B vergleichbar, so gilt also

b) $A \leq \cdot B \supset A \leq \cdot A + B \leq \cdot B$.

Das *starke Mittelwertprinzip* besagt

$A < \cdot B \supset (\neg (A \leq \cdot A + B) \wedge \neg (A + B \leq \cdot A) \vee A < \cdot A + B) \wedge (\neg (B \leq \cdot A + B) \wedge \neg (A + B \leq \cdot B) \vee A + B < \cdot B)$, bzw. im Fall der Vergleichbarkeit von $A + B$ mit A und B

$A < \cdot B \supset A < \cdot A + B < \cdot B$.

[5] Diese Bedingung ist äquivalent mit $A \leq \cdot B \supset (\neg (A + B \leq \cdot A) \vee A \leq \cdot A + B) \wedge (\neg (B \leq \cdot A + B) \vee A + B \leq \cdot B)$.

Zum schwachen Mittelwertprinzip führt folgende intuitive Überlegung: Stehen zwei disjunkte Handlungen A und B zur Wahl, wobei A höchstens so gut ist wie B, so ist A, sofern A mit $A + B$ vergleichbar ist, höchstens so gut wie $A + B$, weil durch $A + B$ im schlechtesten Fall A realisiert wird, während $A + B$ zudem die Möglichkeit eröffnet, daß das evtl. bessere B realisiert wird; ist $A = \cdot B$, so folgt im Falle der Vergleichbarkeit von $A + B$ mit A und B aus 2.5-b $A = \cdot A + B = \cdot B$. Sind andererseits $A + B$ und B vergleichbar, so ist $A + B$ höchstens so gut wie B, weil B nur im günstigsten Fall realisiert wird, während im ungünstigen Fall durch $A + B$ auch A realisiert werden kann.

Diese Überlegung legt es nahe, auch das starke Mittelwertprinzip zu akzeptieren. Dieses ist jedoch nicht unbeschränkt gültig, da $A + B$ ebenso hoch wie B eingestuft wird, falls das Eintreten von A als praktisch ausgeschlossen angesehen wird, und ebenso hoch wie A, falls das Eintreten von B als praktisch ausgeschlossen gilt. Solche Wahrscheinlichkeitsbetrachtungen lassen sich jedoch im gegenwärtig betrachteten Fall nicht präzisieren, wo nur eine Präferenzrelation für die Welten aus I vorliegt, und daher begnügen wir uns zunächst mit dem schwachen Mittelwertprinzip.

Es gibt nun eine Reihe von Festlegungen der Präferenzrelation $\leq \cdot$ für Propositionen mit Hilfe der Relation $\cdot \leq$, die alle den beiden oben genannten Bedingungen genügen und die alle eine gewisse intuitive Plausibilität für sich in Anspruch nehmen können[6]. Wir greifen einige naheliegende Definitionen heraus, wobei wir der intuitiven Durchsichtigkeit wegen zunächst die Relationen $< \cdot$ und $= \cdot$ angeben:

2.6 a) $A <. B := \bigwedge j\,(j \in [B] \supset \bigwedge i\,(i \in [A] \supset i .< j)) \wedge \bigvee j\,(j \in [B]) \wedge \bigvee i\,(i \in [A])$[7]

b) $A <. B := \bigvee j\,(j \in [B] \wedge \bigwedge i\,(i \in [A] \supset i .< j)) \wedge \bigvee i\,(i \in [A] \wedge \bigwedge j\,(j \in [B] \supset i .< j))$

c) $A <. B := \bigvee j\,(j \in [B] \wedge \bigwedge i\,(i \in [A] \supset i .< j)) \wedge \bigwedge j\,(j \in [B] \supset \bigvee i\,(i \in [A] \wedge i .\leq j))$

[6] Vgl. dazu die Erörterung solcher Relationen in Danielsson (68).

[7] Während 2.6-a einen Präferenzbegriff festlegt, der kein Risiko impliziert, sind die folgenden Begriffe »Risiko-Begriffe«, insofern $A <. B$ nicht ausschließt, daß man mit B nicht tatsächlich schlechter fährt als mit A, da es B-Welten gibt, die schlechter sind als einige A-Welten. V. Wright spricht in (63) von *preferences involving*, bzw. *not involving risk*.

d) $A <. B := \mathrm{V}\, i\, (i \in [A] \wedge \Lambda j\, (j \in [B] \supset i .< j)) \wedge \Lambda i\, (i \in [A] \supset \mathrm{V}\, j\, (j \in [B] \wedge i .\leq j))$

e) $A <. B := \mathrm{V}\, j\, (j \in [B] \wedge \Lambda i\, (i \in [A] \supset i .< j)) \wedge \Lambda j\, (j \in [B] \supset \mathrm{V}\, i\, (i \in [A] \wedge i .\leq j)) \vee \mathrm{V}\, i\, (i \in [A] \wedge \Lambda j\, (j \in [B] \supset i .< j)) \wedge \Lambda i\, (i \in [A] \supset \mathrm{V}\, j\, (j \in [B] \wedge i .\leq j))$

f) $A <. \mathrm{B} := \mathrm{V}\, j\, (j \in [B] \wedge \Lambda i\, (i \in [A] \supset i .< j))$

g) $A <. B := \mathrm{V}\, i\, (i \in [A] \wedge \Lambda j\, (j \in [B] \supset i .< j))$.

Für alle diese Relationen gelten die Postulate 2.3—d, e.

Als Gleichheitsrelationen bieten sich folgende Relationen an:

2.7 a) $A =. B := \Lambda i\, (i \in [A] \supset \mathrm{V}\, j\, (j \in [B] \wedge i =. j)) \wedge \Lambda j\, (j \in [B] \supset \mathrm{V}\, i\, (i \in [A] \wedge i =. j))$

b) $A = \cdot B : = \Lambda i\, (i \in [A] \supset \mathrm{V} j\, (j \in [B] \wedge i \cdot \leq j)) \wedge \Lambda i\, (i \in [A] \supset \mathrm{V} j\, (j \in [B] \wedge j \cdot \leq i)) \wedge \Lambda j\, (j \in [B] \supset \mathrm{V} i\, (i \in [A] \wedge i \cdot \leq j)) \wedge \wedge \Lambda j\, (j \in [B] \supset \mathrm{V}\, i\, (i \in [A] \wedge j .\leq i))$

c) $A =. B := \Lambda i\, (i \in [A] \supset \mathrm{V}\, j\, (j \in [B] \wedge i .\leq j)) \wedge \Lambda j\, (j \in [B] \supset \mathrm{V}\, i\, (i \in (A) \wedge j .\leq i))$

d) $A =. B := \Lambda i\, (i \in [A] \supset \mathrm{V}\, j\, (j \in [B] \wedge j .\leq i)) \wedge \Lambda j\, (j \in [B] \supset \mathrm{V}\, i\, (i \in [A] \wedge i .\leq j))$.

Alle vier Relationen genügen den Bedingungen 2.3-a, b, c. Die Relationen (a), (b) erfüllen ferner für jede Relation $< \cdot$ nach 2.6. die Bedingungen 2.3-f, g, während die Relation (c) diese Bedingungen nur für 2.6-f, und die Relation (d) sie nur für 2.6-g erfüllen.

Da die liberalere Festlegung 2.7-b zu 2.6-a bis e paßt, und da in 2.6-f, bzw. 2.6-g nur die jeweils besten, bzw. schlechtesten Welten berücksichtigt werden und dazu die Festlegungen 2.7-c, bzw. 2.7-d passen, definieren wir die Relationen $\leq \cdot$ nach 2.4 zu 2.6-a bis e mit 2.7-b, die Relation $\leq \cdot$ zu 2.6-f mit 2.7-c und die Relation $\leq \cdot$ zu 2.6-g mit 2.7-d.

Wir erhalten dann insbesondere zu 2.6-e und 2.6-f die Relationen

2.8 a) $A \leq. B := \Lambda i\, (i \in [A] \supset \mathrm{V}\, j\, (j \in [B] \wedge i .\leq j)) \wedge \Lambda j\, (j \in [B] \supset \mathrm{V}\, i\, (i \in [A] \wedge i .\leq j))$

b) $A \leq. B := \Lambda i\, (i \in [A] \supset \mathrm{V}\, j\, (j \in [B] \wedge i .\leq j))$[8].

[8] Der Relation nach 2.8-a entspricht die Relation A_4, der Relation nach 2.8-b die Relation A_6 in Danielsson (68).

Alle Relationen $\leq \cdot$ erfüllen das schwache Mittelwertprinzip. Generell gilt $A = \cdot B \supset A = \cdot A + B = \cdot B$. Ist $A < \cdot B$ nach 2.6-a, b, so ist $A + B$ weder mit A noch mit B vergleichbar, d.h. das Mittelwertprinzip gilt in diesem Fall nur im trivialen Sinn. Im Fall 2.6-c gilt $A < \cdot B \supset A < \cdot A + B$, während $A + B = \cdot B$ gilt, falls $A + B$ und B vergleichbar sind. Im Fall 2.6-d gilt $A < \cdot B \supset A + B < \cdot B$, während $A = \cdot A + B$ gilt, falls $A + B$ und A vergleichbar sind. Im Fall 2.6-e gilt entsprechend $B \supset A < \cdot A + B \vee A + B < \cdot B$, es gilt aber generell das Prinzip $A \leq \cdot B \supset A \leq \cdot A + B \leq \cdot B$; d.h. im Fall $A \leq \cdot B$ ist $A + B$ immer sowohl mit A wie mit B vergleichbar. In den Fällen 2.6-f, bzw. 2.6-g gilt darüber hinaus $A \leq \cdot B \supset A + B = \cdot B$, bzw. $A \leq \cdot B \supset A = \cdot A + B$.

Die Relationen $\leq \cdot$ zu 2.6-f und g stellen auch totale Quasiordnungen dar, wenn $\cdot \leq$ eine totale Quasiordnung darstellt, während die übrigen Relationen unabhängig davon nur partielle Quasiordnungen sind.

Nach diesen formalen Überlegungen einige Anmerkungen zur intuitiven Rechtfertigung der Definitionen:

Notwendige Gründe dafür, daß die Proposition B der Proposition A vorzuziehen ist, sind

1. Es gibt B-Welten, die allen A-Welten vorzuziehen sind.
2. Es gibt A-Welten, denen alle B-Welten vorzuziehen sind.

Gibt es keine B-Welten, die allen A-Welten vorzuziehen sind, und auch keine A-Welten, denen alle B-Welten vorzuziehen sind, so ist kein Grund ersichtlich, B dem A vorzuziehen.

Aus den notwendigen Gründen werden hinreichende, wenn wir sie disjunktiv verknüpfen und fordern, daß nicht auch A dem B vorzuziehen ist. Das führt zur Definition 2.6-e, die in diesem Sinn intuitiv am besten fundiert ist. Die Relationen nach 2.6-c, d, f, g berücksichtigen für die Präferenzrelation demgegenüber in asymmetrischer Weise vor allem oder nur die jeweils besten, bzw. jeweils schlechtesten Welten. Die in diesem Sinn symmetrischen Relationen 2.6-a, b sind sehr restriktiv, d.h. sehr viele Propositionen sind danach unvergleichbar, so daß sie zu wenig informativ sind. Die »symmetrische« Relation zu 2.6-c, d ist die Relation 2.6-e. Die »symmetrische« Relation zu 2.6-f, g, die sich durch disjunktive Verknüpfung der definierenden Bedingungen in diesen Definitionen ergibt, führt zu einer trivialen Relation, nach der für alle A und B gilt $A = \cdot B$, falls $\cdot \leq$ eine totale Quasiordnung ist.

Dennoch sei betont, daß nicht etwa nur die Relation $\leq \cdot$ nach 2.8-a (die 2.6-e entspricht) als »vernünftige« Präferenzrelation für Propositionen angesprochen werden kann. Vielmehr erfaßt jede der besprochenen Definitionen einen vernünftigen Sinn von normativer Präferenz, und die Frage einer Auswahl ist daher nicht so sehr eine Frage der Vernünftigkeit oder der generellen Adäquatheit, als eine Frage des Zwecks, den man verfolgt. Wir werden unten sehen, daß man auch gegen die Relation 2.8-a, wie gegen alle allein auf der Basis einer Präferenzrelation $\cdot \leq$ für Welten definierten Relationen $\leq \cdot$ für Propositionen starke intuitive Einwände geltend machen kann, d.h. daß sie alle nur einen sehr speziellen Sinn von normativer Präferenz erfassen. Und wir wollen uns jetzt überlegen, daß z.B. auch die Relation nach 2.8-b eine »vernünftige« Präferenzrelation darstellt:

Der Haupteinwand gegen die Definition 2.8-b ist, daß beim Vergleich von A und B nur die jeweils (im Sinne der Relation $\cdot \leq$) optimalen Welten berücksichtigt werden, so daß generell gilt $A \leq \cdot B \supset A \vee B = \cdot B$. D.h. das starke Mittelwertprinzip ist generell ungültig, und die Alternative zwischen der Proposition, daß eine Person a stiehlt oder ihre Steuerschuld begleicht, wird ebenso hoch eingestuft wie die Proposition, daß die Person a ihre Steuerschuld begleicht. Die tautologische Proposition, daß a stiehlt oder nicht stiehlt, wird maximal bewertet. All das klingt in der Tat sehr unplausibel. Man kann aber diesem Einwand ebenso begegnen wie der Paradoxie von Ross, die sich für die Standardsysteme der deontischen Logik ergibt: Dort gilt $O(A) \supset O(A \vee B)$, eine Folge des Prinzips $A \supset B \vdash O(A) \supset O(B)$: Wenn B logisch aus A folgt und A ist geboten, dann ist auch B geboten. Ist es also geboten, daß a seine Steuerschuld bezahlt, so ist es auch geboten, daß a stiehlt oder seine Steuerschuld bezahlt. In diesem Fall genügt es, darauf hinzuweisen, daß mit B zwar das Gebot $O(A \vee B)$ erfüllt wird, nicht aber evtl. weitere Gebote des Normensystems, das insbesondere auch das Gebot $O(\neg B)$ enthalten kann. Gilt in einem Normensystem $O(A \vee B)$, so verhält man sich also nicht schon damit normengerecht, daß man B tut[9].

[9] Vgl. dazu NWE, S. 20.

Ähnlich kann man für den Präferenzbegriff nach 2.8-b so argumentieren: Wenn aus $A \leq \cdot B$ folgt $A \vee B = \cdot B$, so beinhaltet das nur, daß $A \vee B$ hoch eingestuft wird, insofern B hoch eingestuft wird (folgt $O(A \vee B)$ aus $O(A)$, so ist damit $A \vee B$ nur geboten insofern A geboten ist). Daraus folgt nicht eine hohe Einstufung von A selbst[10].

3 Präferenzgrade und Wahrscheinlichkeiten

Wenn die oben definierten Präferenzrelationen für Propositionen alle sehr spezielle Präferenzbegriffe darstellen, so ist zu fragen, ob man nicht zu einem allgemeineren Präferenzbegriff gelangt, wenn man an die Relation $\cdot \leq$ auf der Menge I der möglichen Welten stärkere Anforderungen stellt. Insbesondere ist zu fragen, ob man nicht den Welten $i \in I$ Zahlen $u(i)$ als Präferenzgrade zuordnen kann, um dann damit den Präferenzgrad $U(A)$ einer Proposition als Mittel aus den $u(i)$ mit $i \in [A]$ zu bestimmen. Damit würde es möglich, in einer intuitiv angemesseneren Weise auch eine Proposition A, für die die meisten Welten aus $[A]$ besser sind als die meisten Welten aus $[B]$, der Proposition B vorzuziehen, selbst wenn es einzelne Welten aus $[A]$ gibt, die schlechter sind als alle Welten aus $[B]$, und einzelne Welten aus $[B]$, die besser sind als alle Welten aus $[A]$.

[10] In van Fraassen (72) und Lewis (74) werden auch indirekte Präferenzen für Welten diskutiert: Es sei Ω eine Menge von Werten, auf der eine Relation $\leq\!*$ definiert ist, die eine (totale) Quasiordnung darstellt. Ist $F(i)$ die Menge der in der Welt i realisierten Werte aus Ω, so kann man in Analogie zu 2.6 in verschiedener Weise Präferenzrelationen für Welten definieren. In Entsprechung zu 2.6-f, bzw. 2.8-b kann man z. B. setzen $i \cdot \leq j := \Lambda x(x \in F(i) \supset Vy(y \in F(j) \wedge x \leq\!* y))$, wo Ω der Wertbereich der Variablen »x«, »y« ist. Man kann aber auch ohne den Umweg über $\cdot \leq$ die Relation $\leq \cdot$ direkt mit Hilfe von $\leq\!*$ definieren, indem man z. B. setzt: $A \leq \cdot B := \Lambda xi(x \in F(i) \wedge i \in [A] \supset Vyj(y \in F(j) \wedge j \in [B] \wedge x \leq\!* y))$. Mit der Relation $\leq\!*$ kann man ersichtlich noch mehr, wenn auch nicht interessantere Präferenzrelationen für Propositionen definieren als mit der Relation $. \leq$. Diese Verallgemeinerung macht für die Logik der Präferenzen und bedingten Obligationen (vgl. D4 und D3 im Abschnitt 6) jedoch keinen Unterschied, so daß wir hier nicht darauf eingehen.

Die Beziehungen zwischen indirekten Präferenzen und Präferenzen für Welten, sowie zwischen diesen und anderen semantischen Strukturen (*Auswahlfunktionen* und *nestings*) wird in Hansson (68) und Lewis (73) und (74) angegeben.

Wenn wir die Präferenzrelation $\leq \cdot$ definieren wollen durch $A \leq \cdot B := U(A) \leq U(B)$, wobei die $U(A)$ Mittel aus den $u(i)$ mit $i \in [A]$ darstellen, so muß die Metrisierung der auf I vorgegebenen komparativen Präferenzrelation eindeutig sein bis auf lineare Transformationen; andernfalls würde die Relation $\leq \cdot$ durch $\cdot \leq$ nicht eindeutig bestimmt. Die Metrisierung einer Quasiordnung $\cdot \leq$ ist aber nur bis auf monoton wachsende Transformationen eindeutig, d.h. mit $u(i)$ ist auch $h(u(i))$ eine Metrisierung von $i \cdot \leq j$, wenn gilt $x < y \supset h(x) < h(y)$[11].

Man wird daher von Quasiordnungen zu sog. *unendlichen Differenzsystemen* übergehen. Ein solches System liegt vor, wenn wir anstelle der 2-stelligen Relation »i wird höchtens so gut bewertet wie j« auf I eine 4-stellige Relation »i wird dem j höchstens so stark vorgezogen wie k dem l« – symbolisch $i, j \cdot \leq k, l$ – definieren, die folgende Eigenschaften hat:

3.1 a) $i, j . \leq k, l \vee k, l . \leq i, j$[12]

b) $i, j . \leq k, l \wedge k, l . \leq m, n \supset i, j . \leq m, n$

c) $i, j . \leq k, l \supset i, k . \leq j, l$

d) $i, j . \leq k, l \supset l, k . \leq j, i$.

Um die Metrisierbarkeit der Struktur sicherzustellen, genügt es nach P. Suppes und M. Winet (55) weiterhin, zu fordern:

e) $\mathrm{V} j\, (i, j . = j, k)$

f) $i . < j \wedge k, l < . j, i \supset \mathrm{V} m\, (i . < m . < j \wedge k, l . \leq j, m)$

g) $i . < j \wedge j, i < . k, l \supset \mathrm{V} mnr\, (k, m . =_r n, l \wedge k, m . \leq j, i)$.

Dabei seien $i, j, k, l, m, n \in I$; r sei eine natürliche Zahl ≥ 1. Die 2-stellige Relation $i \cdot \leq j$ (i wird dem j nicht vorgezogen[13]) wird definiert durch $i, i \cdot \leq j, i$. Man setzt $i \cdot \leq j := \neg (j \cdot \leq i)$, $i \cdot = j := (i \cdot \leq j) \wedge (j \cdot \leq i)$ und entsprechend für die 4-stelligen Relationen $\cdot \leq$ und $\cdot =$, und definiert:

$i, j . =_1 k, l := i, j . = k, l \wedge j . = k$

$i, j . =_{r+1} k, l := \mathrm{V} mn\, (i, j . =_r m, n \wedge m, n . =_1 k, l)$.

[11] Vgl. dazu z.B. Kutschera (72), S. 34. Dort wird unter »Quasireihe« eine totale Quasiordnung im Sinne von 2.1 verstanden.

[12] Hier wird also im Hinblick auf die angezielte Definition $A \leq \cdot B := U(A) \leq U(B)$ die Vergleichbarkeit aller Paare i,j gefordert, da auch die zahlentheoretische Relation $\leq$ konnex ist.

[13] Aus typographischen Gründen bezeichnen wir die 2- und die 4stellige Relation mit dem gleichen Symbol.

Wie Suppes und Winet bewiesen haben, gibt es zu jeder solchen Relation eine Funktion u auf I, deren Werte reelle Zahlen sind, und für die gilt $i \cdot \leq j \equiv u(i) \leq u(j)$ und $i, j \cdot \leq k, l \equiv$ $\equiv u(i) - u(j) \leq u(k) - u(l)$. u ist bis auf lineare Transformation eindeutig, d.h. ist c eine positive reelle Zahl, so ist mit $u(i)$ auch $c \cdot u(i) + d$ eine Metrisierung der Struktur, und zwei Metrisierungen der Struktur gehen auseinander immer durch solche Transformationen hervor[14].

Dieser Hinweis zeigt, wie man von einer komparativen Präferenzrelation auf I zu einer für die Zwecke der Mittelwertbildung hinreichend genau bestimmten metrischen Funktion u auf I gelangen kann.

Die Menge aller möglichen Welten ist von der Mächtigkeit des Kontinuums, wenn wir wie in Δ abzählbar unendlich viele Satzkonstanten annehmen, und jede Menge $[A] \neq \Lambda$ enthält dann überabzählbar viele Elemente. Ein *arithmetisches Mittel* der $u(i)$ ist jedoch nur für abzählbar viele i definierbar[15]. Man wird also $U(A)$ als *gewichtetes Mittel* der $u(i)$ mit $i \in [A]$ bestimmen. Dabei bietet es sich an, Wahrscheinlichkeiten als Gewichte zu wählen. Für die folgenden Überlegungen spielt es keine Rolle, ob wir einen subjektiven oder einen objektiven Wahrscheinlichkeitsbegriff verwenden. Da es hier nicht um subjektive Wertungen, sondern um normative Präferenzen geht, denen eine intersubjektive Geltung eignen soll, wird man jedenfalls nicht an eine subjektive Wahrscheinlichkeit denken, welche die Erwartungen einer einzelnen Bezugsperson charakterisiert, sondern an eine Wahrscheinlichkeit, die allgemeine Annahmen der Gruppe ausdrückt, an die die Normen adressiert sind.

Es sei $\mathfrak{F}$ der Mengenkörper über der Menge I aller möglichen Welten, der genau die Mengen $[A]$ enthält, wo A ein Satz von Δ_0 ist. Auf $\mathfrak{F}$ sei ein Wahrscheinlichkeitsmaß w definiert, sowie für alle $X \in \mathfrak{F}$ mit $w(X) > 0$ eine Funktion U, für die gilt

$$\textbf{3.2} \quad U(A + B) = \frac{U(A) \cdot w(A) + U(B) \cdot w(B)}{w(A) + w(B)}.$$

Statt $w([A])$ schreiben wir dabei kurz $w(A)$ und ebenso für U.

[14] Vgl. dazu auch Kutschera (72), S. 39ff.

[15] Mit einem arithmetischen Mittel bei Beschränkung auf endlich viele Welten arbeitet Rescher in (67).

Der Einfachheit halber wollen wir annehmen, daß für alle nichtkontradiktorischen Sätze A $w(A) > 0$ ist.

Eine solche Funktion erhält man ausgehend von einer Funktion $u(i)$ auf I, für die $\int\limits_I udw$ definiert ist, indem man setzt

3.3 $U(A) = \frac{1}{w(A)} \int\limits_A udw$[16].

Wir können dann setzen:

3.4 $A \leq . B := U(A) \leq U(B)$ für nichtkontradiktorische A und B.

Wenn wir auch $U(A)$ für kontradiktorisches A wieder irgendeinen Wert zuordnen, so ist $\leq \cdot$ eine totale Quasiordnung, da die zahlentheoretische Relation diese Eigenschaft hat. Und es gilt das Mittelwertprinzip in der Form:

3.5 a) $A = . B \supset A = . A + B$

b) $A < . B \supset A < . A + B$ für $w(B) > O$

c) $A < . B \supset A + B < . B$ für $w(A) > O$.

Es stellt sich nun die Frage, wie sich derartige Präferenzbegriffe axiomatisch charakterisieren lassen.

Es sei $\mathbf{D_1}$ das System folgender Axiome und Regeln:

A1: T (T sei eine aussagenlogische Tautologie)

A2: $K(A)$ für kontradiktorische A

A3: $\neg K(A)$ für nicht kontradiktorische A[17].

A4: $A \leq . B \vee B \leq . A$

A5: $A \leq . B \wedge B \leq . C \supset A \leq . C$

A6: $\neg K(A) \wedge A = . B \supset A = . A + B$

A7: $\neg K(B) \wedge A < . B \supset A < . A + B$

A8: $\neg K(A) \wedge A < . B \supset A + B < . B$

A9: $A = . B \wedge \neg (C = . A) \wedge \neg K(C) \wedge A + C = . B + C \wedge \neg K(D) \supset A + D = . B + D$.

[16] Sind umgekehrt U und w gegeben, so erhält man $u(i)$ aus
$$u = \frac{d(U \cdot w)}{dw}.$$

[17] Die Mengen der kontradiktorischen und der nicht kontradiktorischen Sätze der Aussagenlogik sind entscheidbar. $N(A)$ drückt die Eigenschaft $w(A) = 0$ aus.

R1: $A, A \supset B \vdash B$
R2: $A \equiv B \vdash (A \leq\!\cdot\, C) \equiv (B \leq\!\cdot\, C)$
R3: $A \equiv B \vdash (C \leq\!\cdot\, A) \equiv (C \leq\!\cdot\, B)$.

Man überlegt sich leicht, daß für jedes Paar $\langle w, U\rangle$ von Funktionen auf $\mathfrak{F}$, bzw. $\mathfrak{F}-\{\Lambda\}$, so daß w ein Wahrscheinlichkeitsmaß ist mit $w(A) = 0 \equiv [A] = \Lambda$ und U der Bedingung 3.2 genügt – wir bezeichnen dann $\langle U, w\rangle$ als *P-Struktur* – die zugehörige Relation $\leq\cdot$ nach 3.4 diese Bedingungen erfüllt.

Die Umkehrung, daß es zu jeder Relation $\leq\cdot$, die alle Theoreme von $D1$ erfüllt, eine P-Struktur gibt, für die 3.4 gilt, d.h. die Vollständigkeit von $D1$, können wir dagegen nicht behaupten. Nach dem Satz von E. Bolker in (66) und (67) gilt:

Ist $\mathfrak{F}$ ein atomfreier σ-Körper und $\leq\cdot$ eine Präferenzrelation auf $\mathfrak{F}$, die $A2$ bis $A9$ genügt, sowie dem Prinzip
T) Ist $A_1, A_2, \ldots$ eine Folge von Mengen aus $\mathfrak{F}$ mit $A_i \subset A_{i+1}$ (bzw. $A_{i+1} \subset A_i$) und $\neg\, K\,(A_i)$ für alle $i = 1, 2, \ldots$, und gilt für $A = \bigcup_i A_i$ (bzw. $A = \bigcap_i A_i$) $B <\cdot\, A <\cdot\, C$, so gibt es ein n, so daß für alle $m \geq n$ gilt $B <\cdot \bigcup_{i=1}^{m} A_i <\cdot\, C$ (bzw. $B <\cdot \bigcap_i^{m} A_i <\cdot\, C$), so gibt es auf $\mathfrak{F}$ ein Wahrscheinlichkeitsmaß mit $w(A) = 0 \equiv A = \Lambda$ und auf $\mathfrak{F}-\{\Lambda\}$ eine Funktion U, die 3.2 und 3.4 genügt.

Diese Bedingungen für die Existenz einer P-Struktur lassen sich in der Sprache der Aussagenlogik nicht formulieren. Man müßte zu einer Sprache der Aussagenlogik mit unendlich langen Formeln übergehen, in der man auch unendliche Disjunktionen und Konjunktionen bilden kann, um (T) ausdrücken zu können. Darüber hinaus ist aber auch die Existenzbehauptung von Bolker bisher noch nicht vollständig bewiesen worden.

Wir wollen damit unsere Übersicht über die verschiedenen Möglichkeiten beschließen, eine Präferenzrelation $\leq\cdot$ für Propositionen einzuführen. Das Fazit unserer Überlegungen ist: Eine zweistellige Präferenzrelation $\cdot\leq$ für mögliche Welten ist nur eine Basis für die Definition sehr spezieller Präferenzbegriffe für Propositionen. Erst ein metrischer Präferenzbegriff für Welten zusammen mit einem Wahrscheinlichkeitsbegriff bietet eine ausreichende Grundlage für die Definition einer intuitiv befriedigenden und allgemeinen Relation $\leq\cdot$.

Da sich die Eigenschaften einer derartigen Relation bisher jedoch nicht vollständig formalisieren lassen, beschränken wir uns im folgenden auf die Betrachtung spezieller, mit Hilfe von $\cdot \leq$ definierbarer Relationen $\leq \cdot$.

4 Positive und obligatorische Handlungen

Wir wenden uns nun der Frage zu, ob man mit Hilfe der Präferenzrelation $\leq \cdot$ auch Gebote charakterisieren kann. Dazu bieten sich drei Möglichkeiten an:

1. Man zeichnet eine Proposition G als indifferent aus und legt fest $O(A) \equiv G < \cdot A$; d.h. A ist geboten, wenn A einer indifferenten Proposition vorzuziehen ist.

2. Man setzt $O(A) \equiv \neg A < \cdot A$; d.h. A ist geboten, wenn A dem $\neg A$ vorzuziehen ist; wenn es also besser ist A zu tun, als A zu unterlassen.

3. $O(A)$ soll gelten genau dann, wenn es bzgl. $\leq \cdot$ optimale Propositionen B gibt (so daß für alle C gilt $C \leq \cdot B$) und alle diese Propositionen A logisch implizieren.

Welche dieser Möglichkeiten angemessen ist, hängt erstens von der zugrunde gelegten Relation $\leq \cdot$ ab und zweitens von den Eigenschaften, die man für Obligationen fordern will. Im letzten Punkt wollen wir uns darauf festlegen, daß der Operator O die Theoreme des *Standardsystems* **D2** für Gebote erfüllen soll, dem nun die Sprache Δ_2 zugrunde liegt und das durch folgende Axiome und Regeln bestimmt wird:

A1: T

A10: $O(T)$

A11: $O(A \supset B) \wedge O(A) \supset O(B)$

A12: $O(A) \supset \neg O(\neg A)$

R1: $A, A \supset B \vdash B$[18].

Daneben gibt es noch andere Systeme der deontischen Logik, die auch eine gewisse intuitive Plausibilität für sich in Anspruch nehmen können, z.B. weil in ihnen die Paradoxie von Ross oder andere Paradoxien nicht auftreten. Sie sind dann aber entweder so schwach, daß man mit ihnen kaum etwas anfangen kann, oder sie enthalten an anderer Stelle Paradoxien,

[18] Das ist das System D in NWE, S. 46.

die nicht weniger störend sind[19]. Insgesamt kann man wohl sagen, daß das System $D2$ das leistungsfähigste und adäquateste System der deontischen Logik ist, über das wir verfügen[20].

Intuitiv gesehen, erfaßt man durch die Festsetzung $G < \cdot A$ in (1) zunächst die Tatsache, daß A *positiv* oder *gut* ist. Wenn man noch nicht auf spezielle Relationen $\leq \cdot$ Bezug nimmt, können aber auch zwei unverträgliche Propositionen beide positiv sein: Man kann in einem bestimmten Zeitpunkt z.B. entweder die positive Handlung vollziehen, daß man Herrn X im Krankenhaus besucht, oder die positive Handlung, daß man Herrn Y beim Umzug behilflich ist, aber nicht beides zugleich. Da in $D2$ die Regel gilt $A \supset \neg B \vdash O(A) \supset O(\neg B)$, so daß zwei unverträgliche Propositionen nicht beide geboten sein können, kann die Positivität einer Handlung also nicht allgemein als ein hinreichendes Kriterium dafür angesehen werden, daß sie geboten ist. Sie ist auch nicht allgemein ein notwendiges Kriterium für das Gebotensein einer Handlung; denn mit A ist nach dem Prinzip $A \supset B \vdash O(A) \supset O(B)$ auch jede Folge von A geboten; es ist aber nicht gesagt, daß auch jede Folge einer positiven Handlung positiv ist. So kann z.B. die Disjunktion einer positiven mit vielen negativen Handlungen negativ eingestuft werden.

Entsprechendes gilt für die Festsetzung $\neg A < \cdot A$ in (2): die Tatsache, daß eine Handlung A der Unterlassung von A vorzuziehen ist, impliziert noch nicht, daß es geboten ist, A zu tun. Insbesondere kann wieder gelten $\neg A < \cdot A$ und $\neg B < \cdot B$, obwohl A und B unverträglich sind.

[19] Die Paradoxie von Ross tritt z.B. nicht auf, wenn man das Axiom $A10$ ersetzt durch das Axiom $O(A \vee B) \equiv O(A) \wedge O(B)$. Es gilt dann aber $V(A \wedge B) \equiv V(A) \wedge V(B)$, d.h. $A \wedge B$ ist nur dann verboten, wenn sowohl A als auch B verboten sind, und das führt zu dem Theorem $E(A) \supset E(A \wedge B)$. D.h. wenn es erlaubt ist, zu rauchen, so ist es auch erlaubt, zu rauchen und zu stehlen – ein Resultat, das sicher noch weniger akzeptabel ist als die Paradoxie von Ross.

[20] B. van Fraassen hat (73) eingewendet, daß D2 auf die meisten moralischen und juristischen Normsysteme nicht anwendbar sei, da in ihnen Pflichtenkonflikte auftreten, aus denen mit $A12$ folgt, daß in ihnen alles geboten ist. Aus dem gleichen Grund müßte man dann aber die Logik und die subjektive Wahrscheinlichkeitstheorie ablehnen, weil die meisten Systeme von (Wahrscheinlichkeits-) Annahmen, denen wir im Alltag begegnen, nicht konsistent sind.

Für spezielle Relationen kann zwar die Definition $O(A) := G < \cdot A$ oder $O(A) := \neg A < \cdot A$ einen Begriff des Gebotenseins ergeben, der die in $D2$ fixierten Eigenschaften hat. So ist die letztere Definition z.B. akzeptabel für die Relation $\leq \cdot$ nach 2.8-b, sofern diese eine totale Quasiordnung darstellt[21], während sie nach 2.8-a oder nach 3.4 nicht akzeptabel ist. Im allgemeinen erfassen wir aber mit $G < \cdot A$ oder $\neg A < \cdot A$ nicht die Tatsache, daß A geboten ist, sondern nur die Tatsache, daß A positiv ist[22].

Wenn wir wieder von einer Präferenzrelation $i \cdot \leq j$ im Sinne von 2.1 ausgehen, die eine totale Quasiordnung darstellt, und annehmen, daß es bzgl. dieser Ordnung optimale Welten gibt[23], so daß für

4.1 $Q := \{i : \wedge j\, (j \cdot\leq i)\}$,

$Q \neq \Lambda$ ist, so ist A im Sinne von (3) geboten, wenn $Q \subset [A]$ gilt, d.h. wenn A in allen optimalen Welten wahr ist. D.h. wir erhalten

4.2 $O(A) \equiv Q \subset [A]$.[24]

Die Existenz optimaler *Propositionen* als beliebiger Mengen von Welten ist aber weder eine notwendige noch eine hinreichende Bedingung für die Existenz von *Sätzen* A, für die gilt $B \leq \cdot A$ für alle B.

Der Zusammenhang zwischen Geboten und Präferenzen nach 4.2 ist daher nur schwach und indirekt. Insbesondere können wir nicht allgemein Obligationen durch Präferenzen definieren und es gelten nicht allgemein Gesetze wie $O(A) \wedge A$

[21] Vgl. dazu den Abschnitt 6.

[22] Man macht sich leicht klar, daß im Fall (1), wenn man setzt $I(A) \equiv A = \cdot G$ und wegen $O(A) \vee I(A) \vee V(A)$ fordert $V(A) \equiv A < \cdot G$ und die Vergleichbarkeit aller A mit G annimmt, gilt $G < \cdot A \equiv \neg A < \cdot A$, $G = \cdot A \equiv \neg A = \cdot A$ und $A < \cdot G \equiv A < \cdot \neg A$, so daß (1) dann einen Spezialfall von (2) darstellt. – Danielsson definiert in (68) gute, schlechte und indifferente Ereignisse generell durch $A > \cdot T$, $A < \cdot T$ und $A = \cdot T$, wo T eine Tautologie ist.

[23] Die Annahme der Existenz optimaler Welten entspricht der *limit assumption* in D. Lewis (73), 1.4. Diese Annahme erweitert die Klasse der deontologisch wahren Sätze nicht. Vgl. dazu D. Lewis (73), 6.1, sowie den Anhang 7.1.

[24] Vgl. dazu NWE, D 1.9–2.

$\leq \cdot B \supset O(B)$, $I(A) \wedge O(B) \supset A < \cdot B$, $V(A) \wedge O(B) \supset A < \cdot B$, $V(A) \wedge I(B) \supset A < \cdot B$, oder $I(A) \wedge I(B) \supset A = \cdot \cdot B$.[25]

5 Bedingte Gebote

Auch für bedingte Gebote gilt, was wir schon für Präferenzen und Gebote betont haben: Es gibt verschiedene Begriffe bedingten Gebotenseins, die alle in dem Rahmen, in dem sie sich semantisch präzisieren lassen, »vernünftige« Begriffe sind, d.h. eine gewisse intuitive Adäquatheit beanspruchen können.

Für einen *ersten Typ* bedingter Gebote $O(A/B)$ – »Unter der Bedingung B ist A geboten« – gelten z.B. die Prinzipien

*P*1) $B \wedge O(A/B) \supset O(A)$ – falls A unter der Bedingung B geboten ist, und B gilt, so ist A geboten.

*P*2) $O(A/B \vee C) \supset O(A/B)$ – falls A unter der Bedingung $B \vee C$ geboten ist, so ist A auch unter der Bedingung B geboten.

Solche Gebote lassen sich darstellen in der Form $A \supset O(B)$ oder in der Form $A \Rightarrow O(B)$, wo »$\Rightarrow$« eine »wenn–dann« – Beziehung in einem engeren, nicht-extensionalen Sinn darstellt, z.B. eine strikte Implikation. Da jedoch allgemein eine modallogische oder eine intensionslogische Sprache wesentlich ausdrucksreicher ist und erheblich genauere Differenzierungen erlaubt als eine extensionale Sprache, wird ihre Verwendung nicht speziell durch die Darstellung solcher bedingter Obligationen erfordert, so daß wir uns im Rahmen der Aussagenlogik mit der Formulierung $B \supset O(A)$ zufrieden geben können.

Wir wollen im folgenden jedoch einen zweiten Typ bedingter Obligationen untersuchen, wie er z.B. in Danielsson (68), van Fraassen (72) oder Lewis (74) erörtert wird. Danach soll $O(A/B)$ durch einen Präferenzbegriff charakterisiert werden. Man legt zunächst zu jeder Präferenzrelation $\leq \cdot$ eine *bedingte Präferenzrelation* $\leq \cdot_B$ fest durch

[25] Definiert man nach Danielsson (68) gute, schlechte und indifferente Ereignisse durch $G(A) := T < \cdot A$, $S(A) := A < \cdot T$, $I(A) := A = \cdot T$, so gelten für die Relationen nach 2.8-a,b die Beziehungen $G(A) \wedge I(B) \supset A > \cdot B$, $I(A) \wedge S(B) \supset A > \cdot B$, $I(A) \wedge I(B) \supset A = \cdot B$, $I(A) \wedge A < \cdot B \supset G(B)$, $I(B) \wedge A < \cdot B \supset S(A)$.

5.1 $A \leq \cdot_B C := A \wedge B \leq \cdot C \wedge B.$

Für den Vergleich von A mit C unter der Bedingung B kommen also nur die B-Welten in Frage.

$O(A/B)$ soll nun durch $\leq \cdot_B$ in gleicher Weise, d.h. nach den Fällen (1) bis (3), bzw. 4.2 aus Abschnitt 4 definiert werden wie $O(A)$ durch $\leq \cdot$. Je nach dem gewählten Verfahren und nach der zugrundegelegten Präferenzrelation $\leq \cdot$ erhält man dann verschiedene Eigenschaften des Begriffs $O(A/B)$. In jedem Fall gilt aber

5.2 $O(A) \equiv O(A/T)$,

und $O(A/B)$ hat für festes B die Eigenschaften des Begriffs $O(A)$. Geboten ist also, was prima facie, ohne Berücksichtigung der Umstände geboten ist, nicht jedoch, was unter allen Umständen geboten ist.

Es ist nun klar, daß für diesen Begriff einer bedingten Obligation das Abtrennungsprinzip $P1$ nicht gilt: $O(A/B)$ besagt, daß A bei der Präferenzrelation $\leq \cdot_B$ geboten ist, die sich aus $\leq \cdot$ bei Kenntnis von B ergibt, d.h. bei ausschließlicher Berücksichtigung der B-Welten. Ähnlich gilt für bedingte Wahrscheinlichkeiten $w(A/B)$ kein Abtrennungsprinzip der Form $B \wedge w(A/B) = r \supset w(A) = r$, sondern man kann nur sagen: Hat ein Subjekt X eine Wahrscheinlichkeitsbewertung w, für die gilt $w(A/B) = r$, so wird X dem Ereignis A die Wahrscheinlichkeit r zumessen, wenn X weiß, daß B eingetreten ist; m.a.W. wird X bei Kenntnis von B zu einer Wahrscheinlichkeitsbewertung w' mit $w'(A) = w(A/B) = r$ übergehen. Ferner gilt für solche Obligationen auch das Prinzip $P2$ nicht allgemein, denn von $A \leq \cdot_{B \vee C} D$, d.h. von $A \wedge B \vee A \wedge C \leq \cdot D \wedge B \vee D \wedge C$, kann man nur aufgrund spezieller Annahmen über $\leq \cdot$ auf $A \leq \cdot_B D$, d.h. auf $A \wedge B \leq \cdot D \wedge B$ schließen. Aufgrund der folgenden Definition 5.3 gilt $P2$ z.B. nicht.

Nachdem wir uns oben für die Einführung von Obligationen nach 4.2 entschieden haben, wonach die Obligationen nur indirekt, d.h. über die Präferenzrelation $\cdot \leq$ für Welten, mit der Relation $\leq \cdot$ zusammenhängen, und insofern von der Definition von $\leq \cdot$ unabhängig sind, können wir nun eine Formalisierung bedingter Obligationen angeben, die für alle Festlegungen von $\leq \cdot$ passend ist.

Der folgende Interpretationsbegriff stellt sich als eine direkte Verallgemeinerung des Interpretationsbegriffs für unbedingte Gebote dar,[26] wobei wir nun eine Menge S von normativ bewertbaren Welten einführen, um auch Gebote darstellen zu können, die unter *allen* Bedingungen gelten.

5.3 Eine Interpretation von Δ_3 ist ein Quintupel $\langle I, i_0, S, .\leq, \Phi\rangle$ für das gilt:

a) I ist eine Menge von Welten mit $i_0 \in I$.

b) S ist eine Teilmenge von I mit $i_0 \in S$.

c) $.\leq$ ist eine totale Quasiordnung auf I, für die gilt:

  c1) $\neg\, i \in S \supset i .\leq j$ für alle $i, j \in I$.

  c2) $i \in S \wedge \neg\, j \in S \supset j <. \, i$ für alle $i, j \in I$.

  c3) Für alle $X \subset I$ mit $X \cap S \neq \Lambda$ gibt es ein $i \in X$ mit $\bigwedge j \, (j \in X \supset j .\leq i)$.

d) Φ_i ist für alle $i \in I$ eine Funktion, die den Sätzen von Δ_3 Wahrheitswerte (w, f) zuordnet, so daß gilt:

  d1) $\Phi_i(\neg A) = w \equiv \Phi_i(A) = f$

  d2) $\Phi_i(A \supset B) = w \equiv \Phi_i(A) = f \vee \Phi_i(B) = w$

  d3) $\Phi_{i0}(O(A/B)) = w \equiv Q_B \subset [A]$.

Dabei sei $[A] = \{i \colon \Phi_i(A) = w\}$ und $Q_B = \{i \colon i \in [B] \cap S \wedge \bigwedge j \, (j \in [B] \supset j \cdot \leq i)\}$.

5.4 Eine Interpretation $\mathfrak{M} = \langle I, i_0, S, .\leq, \Phi\rangle$ *erfüllt* den Satz A, wenn gilt $\Phi_{i0}(A) = w$. A heißt *allgemeingültig*, wenn A von allen Interpretationen erfüllt wird.

In 5.3 ist i_0 die wirkliche Welt. S ist die Menge der normativ bewertbaren Welten – die Welten aus I–S werden alle als gleichermaßen schlecht eingestuft nach (c1) und (c2). i_0 wird als normativ bewertbar angesehen. Definieren wir

5.5 $N(A) := O(A, \neg A)$,

so gilt $N(A) = S \subset [A]$, sowie $N(A) \supset O(A/B)$ und $N(\neg A) \supset O(B/A)$ für alle B. $N(A)$ besagt also, daß A unter allen Bedingungen geboten ist. Unter normativ unmöglichen Bedingungen A, für die gilt $N(\neg A)$, d.h. $S \cap [A] = \Lambda$ ist dagegen alles geboten; das ist nicht mehr als eine unter formalen Gesichtspunkten praktische, intuitiv aber unerhebliche Festlegung.

[26] Vgl. dazu NWE, D1.9–2.

Es sei nun **D3** der Kalkül, der durch folgende Axiome und Deduktionsregeln bestimmt wird:

A1: T
A13: $O\,(A/A)$
A14: $N\,(A) \supset O\,(A/B)$
A15: $N\,(A \supset B) \wedge O\,(A/C) \supset O\,(B/C)$
A16: $O\,(A/B) \wedge O\,(C/B) \supset O\,(A \wedge C/B)$
A17: $\neg\, O\,(\neg\, B/A) \supset (O\,(C/A \wedge B) \equiv O\,(B \supset C/A))$
A18: $NA \supset A$
R1: $A, A \supset B \vdash B$
R4: $A \vdash N\,(A)$

Es gilt dann der Satz:

5.6 Der Kalkül $D3$ ist vollständig und widerspruchsfrei, d.h. die Theoreme von $D3$ sind genau die allgemeingültigen Sätze.

Aus $D3$ erhält man mit der Definition $O\,(A) :=O\,(A/T)$ sofort das Standardsystem $D2$ der deontischen Logik.

Die semantische Widerspruchsfreiheit von D3 prüft man leicht nach. Die Vollständigkeit von D3 ergibt sich aus der Vollständigkeit des im nächsten Abschnitt angegebenen Systems D4, indem man mit der dort als adäquat ausgewiesenen Definition

5.7 $A \leq\!\cdot\, B := N\,(\neg\, (A \vee B)) \vee \neg\, O\,(\neg\, B/A \vee B)$

die Axiome von D4 als in D3 beweisbar und die Deduktionsregeln von D4 als in D3 zulässig erkennt. Vgl. dazu den Anhang 7.2.

6 Schwache Präferenzen und Gebote

Das Ergebnis unserer Erörterungen war insofern negativ, als 1. eine vollständige Formalisierung des von uns aus intuitiven Gründen bevorzugten Präferenzbegriffs $\leq \cdot$ nach 3.4 im Rahmen der Aussagenlogik nicht gelang, und 2. die von uns aus intuitiven Gründen bevorzugten Begriffe bedingter und unbedingter Obligationen zwar in $D3$ vollständig formalisiert werden konnten, aber keinen direkten und formalisierbaren Zusammenhang zum Präferenzbegriff aufweisen, so daß $D1$ und $D3$ unverbunden nebeneinander stehen.

Daher wollen wir abschließend darauf hinweisen, daß sich der schwache Präferenzbegriff nach 2.8-b im Rahmen der Aussagenlogik vollständig formalisieren läßt und daß sich mit dem schwachen Begriff der Obligation im Sinne der Positivität in diesem Rahmen bedingte und unbedingte Obligationen vollständig charakterisieren lassen. Darin kann man einen formal befriedigenden, intuitiv aber wohl nur bedingt verwendungsfähigen Ersatz für ein umfassendes System von Präferenzen und Obligationen sehen. Vielleicht ergeben künftige Diskussionen jedoch, daß dieses System einen breiteren Anwendungsbereich hat, als es zunächst scheint.

6.1 Eine *Interpretation* von Δ_1 ist eine Interpretation von Δ_3 im Sinne von 5.3, wobei (*d*3) ersetzt wird durch

d3′) $\Phi_{i_0}(A \leq . B) = w \equiv \Lambda i\,(i \in [A] \cap S \supset Vj\,(j \in [B] \cap S \wedge i . \leq j))$.

Die Annahme (*c*3) kann man hier weglassen, sie erweitert aber, wie schon gesagt wurde, nicht die Menge der allgemeingültigen Sätze.

Erfüllungsrelation und Allgemeingültigkeit werden wie in 5.4 definiert.

Wir definieren

6.2 $N(A) := \neg A \leq . K$,

wo K eine Kontraktion ist. Dann gilt wieder $N(A) = S \subset [A]$.

Es sei **D4** der Kalkül, der durch folgende Axiome und Deduktionsregeln bestimmt wird:

A1: T

A4: $A \leq . B \vee B \leq . A$

A5: $A \leq . B \wedge B \leq . C \supset A \leq . C$

A18: $NA \supset A$

A19: $A \leq . B \supset A \vee B = . B$

A20: $N(A \supset B) \supset (A \leq . B)$

R1: $A, A \supset B \vdash B$

R4: $A \vdash NA$

Dann gilt der Satz:

6.3 Der Kalkül $D4$ ist vollständig und widerspruchsfrei, d.h. die Theoreme von $D4$ sind genau die allgemeingültigen Sätze von Δ_1.

Zum Beweis vgl. den Anhang 7.1.

Definiert man nun

6.4 a) $A \leq\cdot_B C := A \wedge B \leq\cdot C \wedge B$

b) $O(A/B) := N(\neg A) \vee \neg A <\cdot_B A$,

so erhält man aus D4 alle Theoreme von D3. Mit

c) $O(A) := O(A/T)$

also auch die Theoreme von $D2$.

Aus 6.4-b ergibt sich

$\Phi_{i_0}(O(A/B)) = w \equiv S \subset [\neg B) \vee Vi(i \in [B] \cap S \wedge Vj(i \,.\!\leq j \wedge j \in [B] \supset j \in [A]))$.

D.h. abgesehen von dem trivialen Fall, daß B normativ unmöglich ist ($S \cap [B] = \Lambda$), gilt $O(A/B)$, wenn es eine B-Welt i aus S gibt, so daß alle B-Welten, die mindestens so gut sind wie i, A-Welten sind. Unter der Bedingung(c3) von 5.3 gilt das genau dann, wenn die besten B-Welten A-Welten sind, d.h. wir erhalten nun, obwohl wir $O(A/B)$ nach 6.4-b durch die Positivität von A unter der Bedingung B definiert haben, die Definitionsbedingung $Q_B \subset [A]$ von 5.3. Bei Zugrundelegung der Präferenzrelation nach 2.8-b läßt sich also $O(A/B)$ durch diese Relation in intuitiv befriedigender Weise definieren.

Im letzten Abschnitt hatten wir die Definition 5.7 : $A \leq \cdot B := N(\neg(A \vee B) \vee \neg O(\neg B/A \vee B)$ benutzt.[27] Man verifiziert nun leicht, daß diese Definition adäquat ist, d.h. daß nach 6.1 und 6.4 gilt $\Phi_{i_0}(O(A/\neg A)) = \Phi_{i_0}(N(A))$ (vgl. 5.5) sowie $\Phi_{i_0}(A \leq \cdot B) = \Phi_{i_0}(N(\neg(A \vee B)) \vee \neg O(\neg B/A \vee B))$.

Wie man leicht beweist, gelten in $D4$ die Sätze $V(A/B) \wedge I(C/B) \supset (A <_B C)$, $I(A/B) \wedge I(C/B) \supset (A =_B C)$, $V(A/B) \wedge O(C/B) \supset (A <_B C) \vee N(\neg B)$, $V(A/B) \wedge (C \leq_B A) \supset V(C/B)$, $I(A/B) \wedge O(C/B) \supset (A =_B C)$, $O(A/B) \wedge (A \leq_B C) \supset E(C/B) \vee N(\neg B)$, $I(A/B) \wedge (A =_B C) \supset E(C/B)$. Statt der letzten drei Prinzipien wären die Sätze $I(A/B) \wedge O(C/B) \supset (A <_B C)$, $O(A/B) \wedge (A \leq_B C) \supset O(C/B)$ und $I(A/B) \wedge (A =_B C) \supset I(C/B)$ intuitiv adäquater. Sie gelten jedoch nicht.

[27] Van Fraassen definiert in (72) $A <\cdot B := O(\neg A/A \vee B)$.

Der Präferenzbegriff für Propositionen nach 2.8-b hat also den Vorteil formaler Einfachheit und den Vorzug, daß sich mit ihm bedingte Obligationen explizit definieren lassen, die auch im Sinne von 5.3 intuitiv adäquat sind.

Der $D4$ zugrundeliegende Präferenzbegriff $\leq \cdot$ nach 2.8-b ist zwar sehr schwach, es würde aber nichts gegen die Verwendung eines schwachen Begriffs sprechen, wenn man damit auch stärkere Begriffe definieren könnte. Zu dieser Frage abschließend zwei Hinweise:

1. Hängen die Werte der Welten i nur davon ab, ob in ihnen die Satzkonstanten $p_1, \ldots, p_n$ gelten, so stellen, wo $E(i)$ derjenige Satz der Gestalt $(\neg)\, p_1 \wedge \ldots \wedge (\neg)\, p_n$ ist, den i erfüllt, die Klassen $[E(i)]$ Äquivalenzklassen bzgl. $\cdot \leq$ dar und es gilt $i \cdot \leq j \equiv E(i) \leq \cdot E(j)$. In diesem Fall, der bei vielen Anwendungen zutrifft, kann man aus $\leq \cdot$ die Relation $\cdot \leq$ rekonstruieren, und kann daher mit $\leq \cdot$ auch alle Präferenzbegriffe nach 2.6 definieren.

2. Auch beim Übergang zu einer ausdrucksreicheren Sprache, wie der Aussagenlogik mit unendlich langen Formeln, kann man die Relation $\cdot \leq$ durch $\leq \cdot$ definieren, da sich dann alle Welten durch Sätze ausdrücken lassen.

Behält man jedoch die Sprache der elementaren Aussagenlogik bei und macht keine Zusatzannahmen wie (1), so lassen sich mit der Relation $\leq \cdot$ nach 6.1 allein keine interessanten Präferenzbegriffe wie z.B. 2.8-a definieren, und $\leq \cdot$ liefert nur sehr schwache Informationen über die Relation $\cdot \leq$.

7 Anhang

7.1 Zum Beweis des Satzes 6.3

Man überzeugt sich leicht davon, daß jede Interpretation nach 6.1 jedes Axiom von $D4$ erfüllt, und daß nach den Deduktionsregeln von $D4$ aus allgemeingültigen Sätzen immer nur allgemeingültige Sätze folgen, daß also $D4$ (semantisch) widerspruchsfrei ist. Daher geben wir hier nur einen Vollständigkeitsbeweis von $D4$ an, der den allgemeinen Grundgedanken für Vollständigkeitsbeweise von L. Henkin folgt[28]. Ein Vollstän-

[28] Vgl. dazu z.B. Kutschera und Breitkopf (71), S. 67ff.

digkeitsbeweis für $D4$ über einer Sprache, die auch iterierte Anwendungen des des Operators $\leq.$ zuläßt, findet sich in D. Lewis (73), Kap. 6. Das Verbot solcher iterierter Anwendungen vereinfacht jedoch den Vollständigkeitsbeweis erheblich.

Wir sagen, daß in einer Ableitung in $D4$ eine Annahmeformel, d.h. eine der Prämissen der Ableitung, von sich selbst *abhängt*, und daß durch die Anwendung der Deduktionsregeln die Abhängigkeit von einer AF sich vererbt.

Wir nennen eine Annahmeformel *kritisch* in einer Ableitung $\mathfrak{H}$, wenn $\mathfrak{H}$ eine Anwendung von $R4$ auf einen in $\mathfrak{H}$ von A abhängigen Satz enthält.

Dann beweist man in der üblichen Weise[29] leicht das *Deduktionstheorem* für $D4$ in der Form:

Gibt es eine Ableitung von B aus $A_1, \ldots, A_n$ in $D4$, in der A_n nicht kritisch ist, so gibt es auch eine Ableitung von $A_n \supset B$ aus $A_1, \ldots, A_{n-1}$ in $D4$, in der nur diejenigen AF kritisch sind, die in der ersten Ableitung kritisch waren.

Schreiben wir $A_1, \ldots, A_n \vdash_\circ B$, falls es eine Ableitung von B aus $A_1, \ldots, A_n$ gibt, in der keine der AF kritisch ist, so folgt aus $A_1, \ldots, A_n \vdash_\circ B$ insbesondere $A_1, \ldots, A_{n-1} \vdash_\circ A_n \supset B$.

Wir benutzen im Beweis neben einigen trivialen Konsequenzen von $A4$ und $A5$ folgende Theoreme von $D4$

T1: a) $A \equiv B \vdash (A \leq. C) \equiv (B \leq. C)$
b) $A \equiv B \vdash (C \leq. A) = (C \leq. B)$.

Beweis: Das folgt aus $A20$ mit $R4$.

T2: $N(A \supset B) \wedge N(A) \supset N(B)$

Beweis: Aus $N(A \supset B)$ folgt $A \wedge \neg B \leq. K$ (a). Nach A20 ist $\neg A \wedge \neg B \leq. \neg A$. Ist $N(A)$, also $\neg A \leq. K$, so ist daher $\neg A \wedge \neg B \leq. K$ (b). Es gilt allgemein $C \leq. D \wedge E \leq. D \supset C \vee E \leq. D$; denn ist $C \leq. E$, so nach A19 $C \vee E =. E \leq. D$, und ist $E \leq. C$, so erhält man ebenso $C \vee E =. C \leq. D$. Aus (a) und (b) erhalten wir damit $A \wedge \neg B \vee \neg A \wedge \neg B \leq. K$, also nach T1a $\neg B \leq. K$, also $N(B)$.

N hat also die Eigenschaften einer Notwendigkeit.

T3: $K \leq. A$

Beweis: Nach R4 gilt $N(K \supset A)$, also nach A20 $K \leq. A$ für alle Sätze A.

[29] Vgl. dazu z.B. Kutschera und Breitkopf (71), S. 62ff.

T4: Gibt es für alle A_i $(i = 1,\ldots, s)$ ein B_k $(k = 1,\ldots, t)$ mit $A_i \leq. B_k$, o gilt $A_1 \vee \ldots \vee A_s \leq. B_1 \vee \ldots \vee B_t$, und umgekehrt.

Beweis: Die Numerierung der $A_1, \ldots, A_s$ und $B_1, \ldots, B_t$ sei so vorgenommen, daß gilt $A_1 \leq. A_2 \leq. \ldots \leq. A_s$ und $B_1 \leq. B_2 \leq. \ldots \leq. B_t$. Nach A19 gilt dann $A_1 \vee \ldots \vee A_s =. A_s$ und $B_1 \vee \ldots \vee B_t =. B_t$. Nach Voraussetzung gibt es zu A_s ein B_k mit $A_s \leq. B_k$, also $A_s \leq. B_t$, also $A_1 \vee \ldots \vee A_s \leq. B_1 \vee \ldots \vee B_t$. Gibt es umgekehrt zu einem i kein k mit $A_i \leq. B_k$, so gilt insbesondere $A_s >. B_t$, also $A_1 \vee \ldots \vee A_s >. B_1 \vee \ldots \vee B_t$.

Wir nennen eine Menge $\mathfrak{A}$ von Sätzen von Δ_1 D4-*konsistent*, wenn es keinen kontradiktorischen Satz K von Δ_1 gibt, so daß in D4 gilt $\mathfrak{A} \vdash_{\circ} K$[30].

Es sei $\Gamma(A)$ die Menge der Satzkonstanten, die im Satz A vorkommen. Dann nennen wir eine Satzmenge $\mathfrak{A}$ $\Gamma(A)$-*D4-maximal*, wenn $\mathfrak{A}$ $D4$-konsistent ist und jede Hinzunahme eines Satzes B mit $\Gamma(B) \subset \Gamma(A)$ zu $\mathfrak{A}$ eine nicht $D4$-konsistente Menge ergibt.

Der Vollständigkeitsbeweis vollzieht sich nun in 2 Schritten:

A) Ist der Satz A nicht in $D4$ beweisbar, so gibt es eine $\Gamma(A)$-$D4$-maximale Satzmenge $\mathfrak{A}$ mit $\neg A \in \mathfrak{A}$.

B) Zu $\mathfrak{A}$ gibt es eine Interpretation $\mathfrak{M} = \langle I, i_0, S, . \leq, \Phi\rangle$, die alle Sätze aus $\mathfrak{A}$ erfüllt.

Damit ist gezeigt: Ist A nicht beweisbar, so ist $\neg A$ erfüllbar, also A nicht allgemeingültig.

Die Behauptung (A) beweist man wie üblich, indem man $\mathfrak{A}$ als Vereinigung der $D4$-konsistenten Mengen $\mathfrak{A}_0, \mathfrak{A}_1, \mathfrak{A}_2, \ldots$ bestimmt mit $\mathfrak{A}_0 = \{\neg A\}$, $\mathfrak{A}_{n+1} = \mathfrak{A}_n \cup \{B_{n+1}\}$, falls $\mathfrak{A}_n \cup \{B_{n+1}\}$ $D4$-konsistent ist, und andernfalls $\mathfrak{A}_{n+1} = \mathfrak{A}_n$; dabei sei $B_1, B_2, \ldots$ eine Abzählung aller Sätze B mit $\Gamma(B) \subset \Gamma(A)$[31].

Die Behauptung (B) beweisen wir wie folgt: Es seien $p_1, \ldots, p_n$ die Satzkonstanten aus $\Gamma(A)$. $q_1, q_2, \ldots$ sei eine Abzählung aller Satzkonstanten von Δ, mit $q_1 = p_1, \ldots, q_n = p_n$. I sei die Menge der Folgen $\langle n_1, n_2, \ldots\rangle$ mit $n_k \in \{0,1\}$ und wir setzen: $\Phi_i(q_m) = w$ genau dann, wenn für $i = \langle n_1, n_2, \ldots\rangle$ $n_m = 1$ ist.

[30] Ist $\mathfrak{A}$ eine unendliche Menge, so soll $\mathfrak{A} \vdash B$ genau dann gelten, wenn es eine endliche Teilmenge $\mathfrak{A}'$ von $\mathfrak{A}$ gibt mit $\mathfrak{A}' \vdash B$.

[31] Vgl. dazu z.B. Kutschera und Breitkopf (71), S. 68f.

$E(i)$ sei derjenige Satz der Form $(\neg)\ p_1 \wedge \ldots \wedge (\neg)\ p_n$, den Φ_i erfüllt.

Wir setzen

1) $S = \{i: (K <. E(i)) \in \mathfrak{A}\}$ und
2) $i .\leq j \equiv (E(i) \leq. E(j)) \in \mathfrak{A}$.

Dann gilt:

3) Für alle Sätze B von Δ, die den Operator $\leq.$ nicht enthalten, ist $\Phi_{i_0}(B) = w \equiv B \in$ A. Das beweist man wie üblich.
4) $i_0 \in S$.
Denn aus $E(i_0) \in \mathfrak{A}$ (nach (3)) folgt mit A18 $K <. E(i_0)$, also $i_0 \in S$ nach (1).
5) $i .\leq j$ stellt wegen A4 und A5 eine totale Quasiordnung auf I dar.
6) $\neg\ i \in S \supset i .\leq j$ für alle $j \in I$.
Aus $\neg\ i \in S$ folgt nach (1) $(E(i) \leq. K) \in \mathfrak{A}$, wegen T1 also $(E(i) .\leq E(j)) \in \mathfrak{A}$, also $i .\leq j$.
7) $\neg\ j \in S \wedge i \in S \supset j <. i$ für alle $i, j \in I$.
Aus $\neg\ j \in S$ folgt wie oben $(E(j)) \leq. E(i)) \in \mathfrak{A}$. Wäre $(E(j) =. E(i)) \in \mathfrak{A}$, so $(E(i) \leq. K) \in \mathfrak{A}$, also $\neg\ i \in S$.
8) $\Phi_{i_0}(A \leq. B) = w \equiv (A \leq. B) \in \mathfrak{A}$.
Ist $\Phi_{i_0}(A \leq. B) = w$, so $\bigwedge i\, (i \in [A] \cap S \supset \bigvee j\, (j \in [B] \cap S \wedge i .\leq j))$. Es sei $E(i_1) \vee \ldots \vee E(i_s)$ die (ausgezeichnete) disjunktive Normalform von A in $p_1, \ldots, p_n$ und ebenso $E(j_1) \vee \ldots \vee E(j_t)$ für B. Dann gilt: Für alle i_m $(m = 1, \ldots, s)$ mit $(K <. E(i_m)) \in \mathfrak{A}$ gibt es ein j_k $(k = 1, \ldots, t)$ mit $(E(i_m) \leq. E(j_k)) \in \mathfrak{A}$. Nach T4 gilt also $(E(i_1) \vee \ldots \vee E(i_s) \leq. E(j_1) \vee \ldots \vee E(j_t)) \in \mathfrak{A}$, also nach T1 $(A \leq. B) \in \mathfrak{A}$. Gibt es kein m mit $(K <. E(i_m)) \in \mathfrak{A}$, so ist $(A \leq. K) \in \mathfrak{A}$ nach A19, also nach T3 und A5 $(A \leq. B) \in \mathfrak{A}$. Die Umkehrung erhält man entsprechend.

Aus (3) und (8) folgt, daß Φ_{i_0} genau jene Sätze B mit $\neg\ (B) \subset \neg\ (A)$ erfüllt, die in $\mathfrak{A}$ enthalten sind, insbesondere also $\neg\ A$.

Der Beweis zeigt, daß die Annahme 5.3-c3, die *limit-assumption* von D. Lewis (vgl. Anmerkung 23) nicht zu zusätzlichen allgemeingültigen Sätzen führt, da sie nach der Konstruktion der Interpretation, die den unbeweisbaren Satz A falsch macht, immer erfüllt wird.

7.2 Zum Beweis des Satzes 5.6

Wir zeigen, daß die Theoreme von $D4$ mithilfe der Definition

5.7: $A \leq. B := N(\neg\ (A \vee B)) \vee \neg\ O(\neg\ B, A \vee B)$

in Theoreme von $D3$ übergehen. [Zur Notation: In *7.2* wird gleichbedeutend $O\ (X, Y)$ statt $O\ (X/Y)$ geschrieben.]

In $D3$ gelten folgende Theoreme:

T1: $N(A \supset B) \wedge N(A) \supset N(B)$.

Beweis: Aus $N(A)$ folgt nach A14 $O(A, \neg B)$, aus $N(A \supset B)$ mit A15 daraus $O(B, \neg B)$, also $N(B)$.

T2: $N(\neg A) \wedge O(A, B) \supset N(\neg B)$

Beweis: Aus $N(\neg A)$ folgt nach A14 $O(\neg A, B)$, mit $O(A, B)$ und A16 folgt daraus $O(A \wedge \neg A, B)$, nach A15 also $O(\neg B, B)$, d.h. $N(\neg B)$.

T3: $\neg N(\neg A) \wedge O(B, A) \supset \neg O(\neg B, A)$

Beweis: Aus $O(B, A)$ und $O(\neg B, A)$ folgt wie unter T2 $O(\neg A, A)$. d.h. $N(\neg A)$.

T4: a) $A \equiv B \vdash O(A, C) \equiv O(B, C)$
b) $A \equiv B \vdash O(C, A) \equiv O(C, B)$

Beweis: (a) folgt direkt aus A15. (b) Es gilt $N(A \supset B) \supset O(B, A)$ nach A13 und A15. Aus $\vdash A \equiv B$ erhält man also $O(A, B)$ und $O(B, A)$. Daraus folgt $O(C, A) \equiv O(C, B)$. Denn aus $O(B, A)$ und $O(C, A)$ folgt nach A 16 $O(B \wedge C, A)$, also für $\neg N(\neg A)$ $O(C, A \wedge B)$; denn aus $O(B \wedge C, A)$ folgt nach A15 $O(B \supset C, A)$, also nach A17 wegen $O(B, A)$, $\neg N(\neg A)$ und T3 $O(C, A \wedge B)$. Nach A17 folgt aus $O(C, A \wedge B)$ mit $O(A, B)$ und T2 $O(A \supset C, B)$, wegen $O(A, B)$ und A16 also $O(C, B)$. Ebenso erhält man aus $O(C, B)$ mit $O(A, B)$ und $O(B, A)$ auch $O(C, A)$. Aus $N(\neg A)$ folgt auch $N(\neg B)$ wegen $O(A, B)$ nach T2. Gilt $N(\neg A)$ und $N(\neg B)$, so gilt $O(C, A)$ und $O(C, B)$, also $O(C, A) \equiv O(C, B)$.

T5: $O(\neg C, B \vee C) \wedge O(A, B \vee C) \supset O(A, B)$

Beweis: Aus $O(B \vee C, B \vee C)$ und $O(\neg C, B \vee C)$ folgt $O(B, B \vee C)$ nach A16, also $O(A, (B \vee C) \wedge B) \equiv O(A, B) \equiv O(B \supset A, B \vee C)$ nach A17 und T2. Aus $O(A, B \vee C)$ folgt $O(B \supset A, B \vee C)$ nach A15, also $O(A, B)$.

T6: $\neg O(\neg A, B) \supset \neg O(\neg (A \vee C), B \vee C)$

Beweis: Aus $O(\neg A \wedge \neg C, B \vee C)$ folgt $O(\neg A, B \vee C)$ nach A15. Daraus folgt nach T3 $O(\neg A, B)$.

Ferner gilt: $N(A) \vee \neg O(T, \neg A) \equiv O(A, \neg A)$, d.h. $N(A) \equiv (\neg A \leq \cdot K)$ nach 6.2 ist äquivalent mit $N(A) \equiv O(A, \neg A)$ nach 5.5. Denn wegen $N(T)$ nach $R4$ gilt nach $A14$ immer $O(T, \neg A)$.

Es ist nun zu zeigen, daß die Axiome nach $A4$, $A5$, $A19$ und $A20$ in $D3$ beweisbar sind.

1) Es gilt $(A \leq \cdot B) \vee (B \leq \cdot A)$, d.h. $N \neg (A \vee B) \vee \neg O(\neg B, A \vee B) \vee \neg O(\neg A, A \vee B)$. Denn aus $O(\neg A, A \vee B) \wedge O(\neg B, A \vee B)$ folgt mit $A16$ $O(\neg A \wedge \neg B, A \vee B)$, also $N(\neg (A \vee B))$.

2) Es gilt $(A \leq \cdot B) \wedge (B \leq \cdot C) \supset (A \leq \cdot C)$, d.h. $(N(\neg (A \vee B)) \vee \neg O(\neg B, A \vee B)) \wedge (N(B \vee C)) \vee \neg O(\neg C, B \vee C)) \supset (N(\neg (A \vee C)) \vee \neg O(\neg C, A \vee C))$.

   Wir erhalten:

   a) $\neg O(\neg B, A \vee B) \wedge \neg O(\neg C, B \vee C) \supset \neg O(\neg C, A \vee C)$. Denn aus $\neg O(\neg B, A \vee B)$ folgt nach T5 $\neg O(\neg (B \vee C), A \vee B \vee C)$, also nach $A17$ $O(B \vee C \supset \neg C, A \vee B \vee C) \equiv O(\neg C, B \vee C)$. Aus $\neg O(\neg C, B \vee C)$ folgt also $\neg O(\neg C, A \vee B \vee C)$ nach T2a (a). Aus $\neg O(\neg C, B \vee C)$ folgt nach T5 ebenso $O(A \vee C \supset \neg C, A \vee B \vee C) \equiv O(\neg C, A \vee C)$, also $\neg O(\neg C, A \vee B \vee C) \equiv \neg O(\neg C, A \vee C)$, wegen ($a$) also $\neg O(\neg C, A \vee C)$.

   b) $N(\neg (A \vee B)) \wedge N(\neg (B \vee C)) \supset N(\neg (A \vee C))$, da N nach $A18$, $T1$, $R4$ die Eigenschaften einer Notwendigkeit hat.

   c) Gilt $N(\neg (A \vee B))$ und $\neg N(\neg (A \vee C))$, so wegen $A14$ $O(\neg A, A \vee C)$. Wäre $O(\neg C, A \vee C)$, so nach $A16$ also $O(\neg (A \vee C), A \vee C)$, d.h. $N(\neg (A \vee C))$, im Widerspruch zur Annahme.

   d) Gilt $N(\neg (B \vee C))$, so $N \neg B$, also nach $A14$ $O(\neg B, A \vee B)$, also im Falle $N(\neg (A \vee B)) \vee \neg O(\neg B, A \vee B)$ $N(\neg (A \vee B))$; man kann dann wie unter (c) argumentieren.

3) Es gilt trivialerweise $(A \leq \cdot B) \supset (A \vee B = \cdot B)$, d.h. $(N(\neg (A \vee B)) \vee \neg O(\neg B, A \vee B)) \supset (N(\neg (A \vee B)) \vee \neg O(\neg B, A \vee B)) \wedge (N(\neg (A \vee B)) \vee \neg O(\neg (A \vee B), A \vee B))$.

4) $N(A \supset B) \supset (A \leq \cdot B)$, d.h. $N(A \supset B) \supset (N(\neg (A \vee B)) \vee \neg O(\neg B, A \vee B))$. Denn aus $N(A \supset B)$ folgt $N(\neg B \supset \neg A)$, also nach $A15$ aus $O(\neg B, A \vee B)$ $O(\neg A, A \vee B)$; und nach $A16$ $O(\neg A \wedge \neg B, A \vee B)$, also $N(\neg (A \vee B))$.

Literaturverzeichnis

Bolker, E. (66): Functions resembling quotients of measures, Transaction of the American Math. Soc. *124* (1966), S. 292–312.

– (67): A simultanous axiomatisation of utility and subjective probability, Philosophy of Science *34* (1967), S. 333–340.

Danielsson, S. (68): Preference and Obligation, Uppsala 1968.

Fraassen, B. van (72): The logic of conditional obligation, Journal of Philosophical Logic *1* (1972), S. 417–438.

– (73): Values and the heart's command, Journal for Philosophy *70* (1973), S. 5–19.

Halldén, S. (57): On the Logic of »Better«, Uppsala 1957.

Hansson, B. (68): Choice structures and preference relations, Synthese *18* (1968), S. 443–458.

Kutschera, F. v. (72): Wissenschaftstheorie, Grundzüge der allgemeinen Methodologie der empirischen Wissenschaften, 2 Bde, München 1972.

– (73): Einführung in die Logik der Normen, Werte und Entscheidungen, Freiburg 1973 (zitiert als »NWE«).

Kutschera, F. v. und Breitkopf, A. (71): Einführung in die moderne Logik, Freiburg 1971.

Lewis, D (73): Counterfactuals, Oxford 1973.

Lewis, D. (74): Semantic analyses for dyadic deontic logics, erscheint in Logical Theory and Semantic Analysis, Essays Dedicated to Stig Kanger on his Fiftieth Birthday, Dordrecht 1974.

Rescher, N. (67): The logic of preference, in: Rescher (Hrsg.): The Logic of Decision and Action, Pittsburgh 1967.

Suppes, P. und Winet, M. (55): An axiomatisation of utility based on utility differences, Management Science *1* (1955), S. 259–270.

Wright, G. H. v. (63): The Logic of Preference, Edinburgh 1963.