FRANZ VON KUTSCHERA

SEMANTIC ANALYSES OF NORMATIVE CONCEPTS

Though substantial progress has recently been made in the semantic analysis of some normative concepts, such as conditional obligation and preference, there is still no systematic survey of the different types of normative notions and no general semantical or logical framework in which these concepts may be analyzed, compared and their relations determined. It is the aim of this paper to contribute to such a systematisation in normative theory.

After a few preliminary remarks the basic types of normative notions are sketched in Section 2. A more detailed discussion and a semantic explication of them is offered in Sections 3–5, and logics for some of them are formulated in Section 6.

1. Preliminaries

If we want to define normative concepts semantically we first have to specify normative languages. They will have to contain normative functors for '... is good', '... is obligatory', '... is to be preferred to –', 'the value of... is –' etc. The languages we shall use are determined first by admitting only sentences as arguments of these functors, i.e. by taking propositions as arguments for the normative concepts that we shall investigate here. In ordinary language we also apply normative concepts to actions and sometimes to objects, but most of these uses may be translated in our way of speaking. Instead of 'For John it is obligatory to pay taxes', for instance, we may say 'It is obligatory, that John pays taxes', instead of 'Paying taxes is obligatory' we may say 'It is obligatory, that all people pay taxes', etc. There is no general method of translation, however, and therefore no general assurance, that all normative sentences may be thus translated.

A second restriction for our normative languages is that we shall base them on a language of propositional logic. This is mainly for the sake of simplicity. Taking a language of predicate logic, moreover, would not raise any essentially new normative problems.

Erkenntnis **9** (1975) 195–218.

The third restriction is also in part motivated by reasons of simplicity: We exclude iterated applications of normative functors and generally sentences of a modal degree higher than one. This, however, is also justified by the fact that the meanings of sentences like 'It is forbidden, that it is allowed, that *A*', 'That *A* is to be preferred to *B*, is to be preferred to *C* having to be preferred to *D*' or 'The value of the value of *A* being *r* is *q*', are very doubtful and it is quite controversial whether such statements make sense at all.[1]

It should also be pointed out that the normative concepts we investigate here are *rational concepts*. They are not intended for the description of what according to existing normative ideas is good, preferable or obligatory, but are based on certain minimal conditions of rationality, e.g. that intensionally equivalent sentences may be substituted for each other in all contexts. Such conditions may also be described as conditions for the consistency or coherence of systems of normative ideas. Similar postulates occur also in deontic logic, for instance, or in the theory of subjective probability. Without such conditions of rationality there would be no logic for normative concepts.

Though we are mainly interested in normative concepts here, it should finally be emphasized, that to normative preferences, values and even to obligations there correspond subjective concepts of preference and value obeying the same logical principles. The concept of subjective preference, that a person, (or a group of persons) prefers *A* to *B*, for instance, has the same logical structure as the normative notion, that (in a normative system) *A* is to be preferred to *B*. So the semantics and logic of the normative concepts may, with another intuitive interpretation of the basic terms, also be employed for their subjective counterparts.

In this paper we shall be almost exclusively concerned with the semantics of normative languages and touch on the question of axiomatizability only briefly in the last section. The important thing in a logical analysis of normative concepts is to have a semantics for them, a calculus is only the next desideratum, and its adequacy (its soundness and completeness) can only be proved by semantic means.

2. Types of Normative Concepts

Normative logic in the comprehensive sense, we think about it here, contains theories of different normative concepts. First there are *classificatory*, *comparative* and *metric* concepts. Since, generally speaking, classificatory concepts can be defined by comparative ones, but not vice versa, and metric concepts have to be introduced on the basis of comparative concepts, the comparative concept of preference suggests itself as the basic notion of normative logic. So we shall first take a look at the different concepts of preference.

1. *Absolute Preference*

The first distinction is that of *absolute* and *relative* concepts or preference. We shall represent the absolute concept by the 2-place relation $A \leqslant .B$ – 'A being the case is not to be preferred to B being the case', or shortly 'A is not to be preferred to B', while the relative concept is represented by the 4-place relation $A, B \leqslant . C, D$ – 'A is not to be preferred more to B than C to D'.

If we consider absolute concepts we have to distinguish secondly unconditional and conditional notions of preference. While a statement $A \leqslant ._C B$ – 'On condition that C A is not to be preferred to B' – about conditional preference refers to a condition C on which A is not to be preferred to B, a statement $A \leqslant .B$ about unconditional preference contains no explicit reference to such conditions.

a. *Unconditional preference.* We shall investigate three notions of (absolute) unconditional preference.

a1. *Intrinsic preference*: If A is intrinsically not to be preferred to or not better than B – we shall express this by $A \leqslant_0 B$ – that means that A as such, considered by itself and aside from what else may be the case, is not better than B. This implies that in evaluating A and B we do not take into account any given circumstances or factual assumptions or expectancies as to what A- and B-worlds are or will be realized, i.e. in what factual contexts A and B are considered. Therefore intrinsic betterness coincides with *prima facie betterness*. It is the fundamental notion of betterness, from which all others derive.

Intrinsic betterness is to be distinguished from betterness in all circumstances, i.e. in all worlds, which may be represented as logical truth of $A \leqslant_0 B$, and from betterness under all conditions $\bigwedge C(A \leqslant_{0C} B)$, where $A \leqslant_{0C} B$ is defined in the manner of (b) below.

a2. *Normal preference*: Often a statement of preference is understood not in the sense of $A \leqslant_0 B$ but as 'That things are the way they normally (usually) are if A is not to be preferred to things being the way they normally are if B' or as 'That things are the way they would be in case of A, is not to be preferred to things being the way they would be in case of B' – in symbols $A \leqslant_s B$. Then we do not consider all and even the remotest possibilities of A or B being the case (all A- and all B-worlds), but only the normal ways in which A and B are realized. If $f(A)$ is the set of normal A-worlds in this sense, so that it represents the propositions that things are the way they normally are in case of A, then $A \leqslant_s B$ is equivalent to $f(A) \leqslant_0 f(B)$.

a3. *Probabilistic preference*: If we evalutate the propositions A and B not just by excluding some ways in which they might come true, but base their evaluation on probabilities assigned to the ways in which they may come true (i.e. to the A- and B-worlds) we obtain still another relation $A \leqslant_p B$ of absolute unconditional preference. $A \leqslant_p B$ then means that the expected value of A is not greater than the expected value of B. If the propositions A and B concern actions then with $A \leqslant_p B$ we compare not the intrinsic value of the actions but the values of their expected outcomes. This is a very different concept from that of intrinsic betterness. A proposition may be intrinsically very good (very good in principle), but its expected outcome (its practical use, considering the likely circumstances) may be almost valueless.

b. *Conditional preference*. If we turn to conditional preference now, to each of the three notions of unconditional preference, $\leqslant_0$, $\leqslant_s$, $\leqslant_p$, there corresponds a conditional form: If the index '.' may stand for indices '$_0$', '$_s$' or '$_p$' then $A \leqslant_{.C} B$ means, that on condition that C A is not to be preferred to (not better than) B in the sense of $\leqslant_.$.

Conditional concepts of absolute preference are therefore 3-place relations. They may be defined from the 2-place relations by $A \leqslant_{.C} B := A \wedge C \leqslant_. B \wedge C$, since on condition that C we only have to consider C-worlds among the A- and B-worlds. If C is a tautology, i.e. a completely uninformative condition, then we have $A \leqslant_{.T} B \equiv A \leqslant_. B$, and that, in case of $\leqslant_0$, justifies our speaking of *prima facie* preference.

2. *Relative Preference*

All the distinctions of the absolute case carry over to the relative case

$A, B \leqslant .C, D$. We therefore have here 4-place relations $\leqslant_0$, $\leqslant_s$ and $\leqslant_p$ and 5-place-relations $\leqslant_{0C}$, $\leqslant_{sC}$, $\leqslant_{pC}$ with the 5-place relations $A, B \leqslant ._C D, E$ definable from the 4-place relations.

We may of course, also derive other preference-relations from the basic ones indicated. $A \wedge C \leqslant .B \wedge D$ may, for instance, also be read as 'A on condition that C is not to be preferred to B on condition that D', and $A \wedge B, B \leqslant .C \wedge D, D$ may be read as 'A on condition that and relatively to B is not to be preferred to C on condition that and relatively to D'.[2] We may define all the absolute concepts from the relative ones by setting $A \leqslant .B := A, B \leqslant .A, A$, but not *vice versa.*

If we go over to *metric concepts* of (normative) value, we can again distinguish absolute statements about value of the form $V(A)=r$: 'The value of A being the case (or shortly 'of A') is r' from relative ones. But since such relative statements have the form $V(A)-V(B)=r$, the relative concepts are definable from the absolute concepts and are thus without any interest of their own.

In correspondence to the relations $\leqslant_0$, $\leqslant_s$, $\leqslant_p$ we can introduce metric values V_0, V_s and V_p so that $V_0(A)$ is the *prima facie* value of A, $V_s(A)$ is value A normally has, and $V_p(A)$ is the expected value of A. Conditional values $V.(A, B)$ are the values (*prima facie*, normal or probabilistic) on condition that B. As in the comparative case they will be equal to $V.(A \wedge B)$.

We can define the comparative concepts from metric ones by setting $A \leqslant .B \equiv V.(A) \leqslant V.(B)$ and $A, B \leqslant .C, D \equiv V.(A) - V.(B) \leqslant V.(C) - V.(D)$, but not *vice versa.* The metric concepts $V.$, however, are to be introduced on the basis of the comparative ones $\leqslant .$ by showing that there are real-valued functions $V.$ obeying these equivalences unique up to a certain class of transformations. Since the 2-place relation of absolute preference will generally have too little structure to make the values unique in an informative way, the concepts of relative preference are better objects for metrisation. If they constitute an algebraic difference system in the sense of Krantz *et al.* (1971), 4.4, then a function $V.$ for which the second equivalence holds exists and is unique up to positive linear transformations.

For the *classificatory concepts* (e.g. $G.(A)$ – 'It is good that A' –, $B(A) \equiv \equiv G(\neg A)$ – 'It is bad that A' –, and $I(A) \equiv \neg G(A) \wedge \neg G(\neg A)$ – 'It is indifferent whether A') the relative concepts seem to be of no interest in the normative case.[3] We have then unconditional intrinsic goodness G_0 and conditional concepts of goodness: goodness in the normal cases G_s, goodness of the expected outcome G_p and goodness of A on condition that B: $G.(A, B)$.

We may introduce a *weak* concept of goodness on the basis of preferences by setting $G.A := \neg A <. A$ (and hence $B.A := A <. \neg A$ and $IA := \neg A = .A$) or $G.(A, B) := \neg A <._B A$ etc.

Therefore we do not generally have $G.(A, B) \equiv G.(A \wedge B)$ here.

We may also define goodness in the *strong* sense – symbolically G^* – as optimality, so that G^*A holds iff there are no better propositions than A. What is *obligatory* derives from the normative concept of optimality by defining a proposition A as obligatory iff all the optimal propositions or all the best worlds are included in A. This comes to saying that it is obligatory to do A iff all ways of not doing A are worse than some ways of doing A. Since in the standard logics of obligation we have $OA \supset O(A \vee B)$ we cannot set $O.A \equiv G.^*A$, for $G.^*A \supset G.^*(A \vee B)$ is generally not true. $G.^*A$ implies $O.A$ (it is obligatory to realize one of the best worlds), but $O.A$ also holds if for some B we have $G.^*B$ and B implies A. For obligations the *prima facie* notions O_0A and $O_0(A, B)$ seem to be of the main interest.

From the classificatory normative concepts we can generally not define the comparative ones. For an exception see the end of the next section.

This survey shows that there are quite a lot of different intuitive normative concepts. When we try now to explicate these notions semantically this fact must be kept in mind. We cannot expect that all notions of preference, goodness or value have the same logical structure, so that, if a logical consequence of our definitions seems to be counter-intuitive, we always have to ask, whether this intuition really is based on the specific notion that was to be captured in the definition.

3. *Intrinsic Value*

The simplest approach to a definition of intrinsic preference between propositions is to base it on a preference-relation for worlds. This may seem foreign to the notion of intrinsic betterness, since in this fashion we compare whole worlds in which a proposition is realized, i.e. we take into account all the circumstances under which it may be true. But the value of some state of affairs always depends on the circumstances and therefore its intrinsic value should be defined by abstraction from its value in concrete situations. To determine the value of concrete situations from that of the propositions realized in it would be much more difficult since the value of A and B as such say nothing about the value of $A \wedge B$; A and B may both be good, for instance, and $A \wedge B$ may still be indifferent or bad.

If I then is a set of worlds, a two-place relation $.\leqslant$ is to be defined on I so that $i.\leqslant j$ means that i is (normatively) not better than j. $.\leqslant$ is to be a weak ordering, so that we have

(3.1) (a) $\quad i.\leqslant j \vee j.\leqslant i$

and

(b) $\quad i.\leqslant j \wedge j.\leqslant k \supset i.\leqslant k.$[4]

As usual we define

(D3.2) (a) $\quad i.<j := \neg(j.\leqslant i)$

(b) $\quad i.=j := i.\leqslant j \wedge j.\leqslant i.$

If we want to determine an intrinsic preference-relation $X.\leqslant Y$ for propositions, i.e. for subsets of I, by $i.\leqslant j$, quite a number of possibilities suggest themselves, from which we mention only three:[5]

(I) (a) $\quad X.\leqslant Y := \bigwedge i(i\in X \supset \bigwedge j(j\in Y \supset i.\leqslant j) \wedge Y \neq \bigwedge)$

(b) $\quad X.\leqslant Y := \bigwedge i(i\in X \supset \bigvee j(j\in Y \wedge i.\leqslant j)) \wedge \bigwedge j(j\in Y \supset \bigvee i(i\in X \wedge$
$\wedge\, i.\leqslant j))$

(c) $\quad X.\leqslant Y := \bigwedge i(i\in X \supset \bigvee j(j\in Y \wedge i.\leqslant j)).$

According to (Ia) Y is at least as good as X iff all the Y-worlds are at least as good as all the X-worlds; i.e. whatever else may happen, as long as Y holds we fare at least as good as if X holds. This is certainly sufficient for intrinsic preference but not necessary: In some Y-world j something really bad may be the case which is not the case in an X-world i. Then $X.\leqslant Y$ would not hold according to (Ia), although $j.<i$ is not due to X or Y. Furthermore (Ia) does not even define a partial weak ordering, since it is not reflexive.

In the sense of (Ib) Y is at least as good as X iff in every case of Y there are X-worlds which would have been at least as bad, and in every case of X there are Y-worlds which would have been at least as good. Though with Y we may actually fare worse than with X, there is always consolation in the fact that with X we might have fared equally bad if not worse, and certainly not better than in more fortunate Y-cases.

According to (Ic), finally, Y is at least as good as X iff X allows no possibilities that are better than all the Y-possibilities. In other words: X is to be preferred to Y iff X allows for better worlds than Y. Is this condition sufficient for intrinsic preference, or should it be complemented, as in (Ib), by the condition, that X does not also allow for worse contingencies than Y? Or could we not with equal right say that X is preferable to Y iff X excludes worse possibilities than Y? These questions cannot be answered simply by 'yes' or 'no'. There are a lot of ways to define the value of a proposition X from the values of the X-worlds. Beside the ones mentioned above we shall discuss in Section 5 also the possibility to define the value of X as an average of the values of the X-worlds, for instance. So our choice will have to be determined by considering how adequate an explanandum of the intuitive notion of intrinsic preference each definition provides and how simple it is. It seems natural and sufficient on the one hand to determine the value of X by the best possibilities X allows. And then there is the formal advantage of (Ic) against (Ib) that it defines a total, not only a partial weak ordering. That we can compare all propositions is, of course, a very strong idealization, but if we assume, as in (3.1), that we can compare all worlds, then we should also be able to compare all propositions. On the other hand some consequences of (Ic) seem to be counterintuitive: From (Ic) we obtain the principle

(II) $\quad X. \leqslant Y \supset X \cup Y. = Y,$

therefore the proposition that I pay taxes or steal is considered intrinsically just as good as that I pay my taxes. This principle corresponds to the deontic principle $OA \supset O(A \vee B)$, which leads to the paradox of Ross. Just as this 'paradox' is eliminated by pointing out that from $O(A \vee B)$ it does not follow that B is obligatory or even permitted, so it must be emphasized however, that from $X \cup Y. = Y$ it does not follow that X is equally good as Y or good in itself. With (Ib) we would not do any better with respect to (II). So if we want to base our semantics for $\leqslant_0$ solely on a preference relation for worlds we cannot do much better but certainly worse than chosing (Ic).

We might also understand the definitions (I) as principles for decision under uncertainty, i.e. in the absence of a probabilistically calculable risk. (Ic) would then certainly be too simple and over-optimistic. But, as we tried to explain in the last section, the relation $\leqslant_0$ is no substitute for

$\leqslant_p$ in case no probabilities are defined, but is to express *intrinsic* value, not any rudimentary sort of expected value.

Let $\mathfrak{P}_1$ now be the language obtained from the language $\mathfrak{A}$ of propositional logic by stipulating that $A \leqslant_0 B$ is to be a sentence of $\mathfrak{P}_1$ if A and B be sentences of $\mathfrak{A}$. Then we obtain the following concept of an interpretation from (Ic):

(D3.3) An interpretation of $\mathfrak{P}_1$ is a quintupel $\langle I, i_0, S, .\leqslant, \Phi\rangle$ such that

(1) I is a set of worlds.

(2) S is a subset of I with $i_0 \in S$.

(3) $.\leqslant$ is a weak ordering on I such that

(a) $\neg i \in S \supset i. \leqslant j$

(b) $\neg i \in S \wedge j \in S \supset i. < j$

(c) $X \neq \Lambda \supset \bigvee i(i \in X \wedge \bigwedge j(j \in X \supset j. \leqslant i))$.

(4) For all $i \in I$ Φ_i is a function from the set of sentences of $\mathfrak{P}_1$ into the set of truth-values $\{t, f\}$ such that

(a) $\Phi_i(\neg A) = t$ iff $\Phi_i(A) = f$, and $\Phi_i(A \wedge B) = t$ iff $\Phi_i(A) = \Phi_i(B) = t$,

(b) $\Phi_i(A \leqslant_0 B) = t$ iff $\bigwedge i(i \in [A] \cap S \supset \bigvee j(j \in [B] \cap S \wedge i. \leqslant j))$.

Here $[A]$ is to be the set $\{i \in I : \Phi_i(A) = t\}$.

(D3.4) An interpretation $\mathfrak{J} = \langle I, i_0, .\leqslant, \Phi\rangle$ *satisfies* the sentence A iff $\Phi_{i_0}(A) = t$. A is $\mathfrak{P}_1$-*valid* iff all interpretations of $\mathfrak{P}_1$ satisfy A.

i_0 is to be the 'real' world and S the set of worlds normatively evaluable from the standpoint of i_0. If we define normative necessity by

(D3.5) $NA := \neg A \leqslant_0 \bot$,

where $\bot$ is a contradiction, we have

(3.6) $\Phi_i(NA) = t$ iff $S \subset [A]$.

If we should be interested also in sentences with iterated applications of the operator $\leqslant_0$, the relation $.\leqslant$ and the set S should have to depend on the world i, since from the standpoint of different worlds different other worlds may be evaluable or preferable to others. The *limit-assumption* (D3.3–3c) could have been left out in (D3.3). It is not plausible in every case that there are optimal worlds instead of infinite sequences of better and better and better worlds. But for the resulting logic this assumption makes no difference, since every satisfiable sentence of a propositional language may be satisfied by an interpretation over a finite set I.[6] And we need this assumption explicitly in the definition of relative preference.

If we set

(D3.7) $g(A) := \{i : i \in [A] \cap S \wedge \bigwedge j(j \in [A] \supset j. \leqslant i)\}$

so that $g(A)$ is the set of optimal $[A] \cap S$-worlds and $g(A) = \bigwedge \equiv \equiv [A] \cap S = \bigwedge$, we obtain from (D3.3–4b)

(3.8) $\Phi_i(A \leqslant_0 B) = t$ iff $\bigwedge i(i \in g(A) \supset \bigvee j(j \in g(B) \wedge i. \leqslant j))$.

Since $\leqslant_0$ is a total weak ordering we may define

(D3.9) (a) $A < .B := \neg(B \leqslant .A)$
(b) $A = .B := (A \leqslant .B) \wedge (B \leqslant .A)$.

Let us turn to *relative prima facie* preference now. $\mathfrak{P}_2$ is to be that language that is obtained from $\mathfrak{A}$ by stipulating that $A, B \leqslant_0 C, D$ be a sentence of $\mathfrak{P}_2$ if A, B, C and D be sentences of $\mathfrak{A}$.

(D3.10) An interpretation of $\mathfrak{P}_2$ is a quintuple $\langle I, i_0, S, . \leqslant, \Phi \rangle$ such that

(1) I is a set of worlds.

(2) S is a proper subset of I with $i_0 \in S$.

(3) $. \leqslant$ is a relation on I with

(a) $i, j. \leqslant k, l \vee k, l. \leqslant i, j$

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

(c) $i, j. \leqslant k, l \supset i, k. \leqslant j, l$

(d) $i, j. \leqslant k, l \supset l, k. \leqslant j, i$

(e) $\neg i \in S \supset i. \leqslant j$

(f) $\neg i \in S \wedge j \in S \supset i. < j$

(g) $X \neq \bigwedge \supset \bigvee i(i \in X \wedge \bigwedge j(j \in X \supset j. \leqslant i))$.

(4) For all $i \in I$ Φ_i is a function from the set of sentences of $\mathfrak{P}_2$ into $\{t, f\}$ such that

(a) $\Phi_i(\neg A) = t$ iff $\Phi_i(A) = f$ and $\Phi_i(A \wedge B) = t$ iff $\Phi_i(A) = \Phi_i(B) = t$

(b) $\Phi_i(A, B \leqslant_0 C, D) = t$ iff $\bigvee ijkl(i \in g'(A) \wedge j \in g'(B) \wedge \wedge k \in g'(C) \wedge l \in g'(D) \wedge i, j. \leqslant k, l)$.

In (D3.10) we make use of the definitions

(D3.11) (a) $i. \leqslant j := i, j. \leqslant i, i$
(b) $g'(A) := \{i \colon i \in [A] \cup \bar{S} \wedge \bigwedge j(j \in [A] \cup \bar{S} \supset j. \leqslant i)\}$.

$i. \leqslant j$ is then a weak ordering as postulated in (3.1).

(D3.10) calls for some explanations again:

(1) According to (3a, b) $i, j. \leqslant k, l$ is a weak ordering for pairs of worlds. (3c, d) are necessary conditions for $. \leqslant$ to be a relation between differences in value. S is again the set of evaluable worlds. In view of (4b) we postulate in (2) that $\bar{S} \neq \bigwedge$, i.e. that there are worst worlds. This, as the limit-assumption, is harmless for propositional languages.

(2) According to the 2-place relation $\leqslant_0$ the empty proposition $\bigwedge$ is equal to $\bar{S}$. If we want to define the 4-place relation $\leqslant_0$ also for empty propositions we have to make them equal again to $\bar{S}$. This is achieved with the help of the function g'.

We define

(D3.12) $A \leqslant_0 B := A, B \leqslant_0 A, A$.

Normative necessity N may be defined as in (D3.5).

From (D3.12) we obtain

(3.13) $\Phi_i(A \leqslant_0 B) = t$ iff $\bigwedge i(i \in g(A) \supset \bigvee j(j \in g(B) \wedge i. \leqslant j))$,

i.e. (3.8).

Questions of metrisation are beyond the scope of this paper. So we shall only say that, if by adding suitable postulates we make the 4-place relation $. \leqslant$ on I into an algebraic difference system, there exists a real-valued function v, unique up to positive transformations, for which we have $v(i) - v(j) \leqslant v(k) - v(l) \equiv i, j. \leqslant k, l$. From this we obtain a function V_0 with

(3.14) $V_0(A) - V_0(B) \leqslant V_0(C) - V_0(D) \equiv A, B \leqslant_0 C, D$,

if we set $V_0(A) = r$ if for some $i \in g'(A)$ we have $v(i) = r$.

If we turn to the *classificatory prima facie concepts* now, weak goodness coincides with strong goodness.

We define

(D3.15) (a) $G_0A := \neg A <_0 A$
(b) $G_0(A, B) := \neg A <_{0B} A$

and obtain

(3.16) (a) $\Phi_i(G_0A) = t$ iff $g(I) \subset [A]$,
(b) $\Phi_i(G_0(A, B)) = t$ iff $g(B) \subset [A] \wedge g(B) \neq \bigwedge$.

That is: A is good in the weak sense iff the best worlds are A-worlds. Then there is no better proposition than $[A]$. And A is good in the weak sense on condition that B iff there are best B-worlds and they all are A-worlds. Then A is also optimal on condition that B. And since in view of (II) we have $G_0A \supset G_0(A \vee B)$, what is good is also obligatory and vice versa. But, as for $N\neg B$ all A are indifferent on condition that B, we may postulate $N\neg B \supset O(A, B)$ for all A in order to obtain the more customary laws for conditional obligations. We define therefore

(D3.17) (a) $OA := G_0A$
(b) $O(A, B) := G_0(A, B) \vee N\neg B$.
(c) $V(A, B) := O(\neg A, B)$
(d) $I(A, B) := \neg O(A, B) \wedge \neg O(\neg A, B)$
(e) $P(A, B) := \neg O(\neg A, B)$.

We have then $\Phi_i(O(A, B)) = t$ iff $g(B) \subset [A]$, $\Phi_i(I(A, B)) = t$ iff $g(B) \cap [A] \neq \bigwedge \wedge g(B) \cap [\neg A] \neq \bigwedge$, and $\Phi_i(V(A, B)) = t$ iff $g(B) \cap [A] = \bigwedge$. Therefore the following connections of deontic concepts with intrinsic preference obtain:

(3.18) (a) $V(A, B) \wedge I(C, B) \supset (A <_{0B} C)$
(b) $I(A, B) \wedge I(C, B) \supset (A =_{0B} C)$
(c) $I(A, B) \wedge O(C, B) \supset (A =_{0B} C)$
(d) $V(A, B) \wedge O(C, B) \supset (A <_{0B} C) \vee N\neg B$.

(a), (b) and (d) are satisfactory, but instead of (c) we would like to have $I(A, B) \wedge O(C, B) \supset A <_{0B} C$, which is not valid however.

We also have

(3.19) (a) $V(A, B) \wedge (C \leqslant_{0B} A) \supset V(C, B)$
(b) $O(A, B) \wedge (A \leqslant_{0B} C) \supset P(C, B) \vee N\neg B$
(c) $I(A, B) \wedge (A =_{0B} C) \supset P(C, B)$.

Instead of (b) and (c) the laws $O(A, B) \wedge A \leqslant_{0B} C \supset O(C, B)$ and $I(A, B) \wedge \wedge A =_{0B} C \supset I(C, B)$ would be intuitively more satisfactory, but they again are not valid.

For obligations (and hence for intrinsic goodness and preference) we can also define interpretations on the basis of the function g:

(D3.20) An interpretation of $\mathfrak{P}_1$ is a quadruple $\langle I, i_0, g, \Phi\rangle$ such that

(1) I is a set of worlds with $i_0 \in I$.

(2) g is a function from subsets of I to subsets of I, so that

(a) $g(X) \subset X$

(b) $X \subset Y \wedge g(X) \neq \Lambda \supset g(Y) \neq \Lambda$

(c) $X \subset Y \wedge g(Y) \cap X \neq \Lambda \supset g(X) = g(Y) \cap X$

(d) $i_0 \in S$, where $S = U_X g(X)$.

(3) For all $i \in I$ Φ_i is a function from the set of sentences of $\mathfrak{P}_1$ into $\{t, f\}$ so that

(a) $\Phi_i(\neg A) = t$ iff $\Phi_i(A) = f$, and $\Phi_i(A \wedge B) = t$ iff $\Phi_i(A) = \Phi_i(B) = t$

(b) $\Phi_i(O(A, B)) = t$ iff $g(B) \subset [A]$.[7]

$g(A)$ is again the set of optimal A-worlds. In the same manner we could define a g'-semantics for relative preference.

In view of $(A \leqslant_0 B) \equiv N\neg(A \vee B) \vee \neg O(\neg B, A \vee B)$ (where NA may now be defined by $O(A, \neg A)$) and $A \leqslant_0 B \equiv \neg G_0(\neg B, A \vee B)$ we may also define $\leqslant_0$ from O or G_0, so that in the case of notions of intrinsic value, as we have determined them, the (absolute) comparative concepts can also be defined by the classificatory concepts.

4. *Normal Value*

We turn now to the second of the three basic normative concepts and begin again with the 2-place preference relation, i.e. with $\leqslant_s$. Here we now have to introduce a selection function f, so that $f(X)$ for $X \subset I$ is the set of worlds possible on condition that X in the sense that in case of X normally one of the $f(X)$ worlds is realized. This *conditional possibility* is to be distinguished from *absolute possibility* which may be defined as possibility under some condition. If we set $R := U_X f(X)$, then R is the set of absolutely possible worlds.

Selection functions of this kind are used in the logic of conditionals. They were first introduced by Stalnaker (1968). A conditional $K(A, B)$

'If it is the case that B, then it is the case that A' is defined to be true if $f(B) \subset [A]$, i.e. if A holds in all normal B-worlds. Stalnaker and Lewis (1973) interpret $f(X)$ as the set of X-worlds most similar to ours. Such an interpretation calls for the condition of *strong centering* $i_0 \in X \supset f(X) = \{i_0\}$, since i_0 is more similar to itself than any other world. This, however, is not adequate for the interpretation of indicative conditionals, since it leads to the principle $A \wedge B \supset K(A, B)$, i.e. all indicative conditionals with true antecedens and succedens would then be true. Therefore we stick to the interpretation of $f(X)$ as the set of normal X-Worlds.[8]

We may bring the notion of normal cases into the definition of preference in a natural way by saying: A being the case is to be preferred to B being the case iff the state of things being the way they normally are if A is to be preferred to the state of things being the way they normally are if B.

If we would, following Åqvist (1973), introduce a functor F into the object language, defined by $\Phi_i(FA) = t$ iff $i \in f(A)$, we could define the relation $\leqslant_s$ by $A \leqslant_s B := FA \leqslant_0 FB$, and $K(A, B) := L(FB \supset A)$, where L is defined by $\Phi_i(LA) = t$ iff $R \subset [A]$. Since we want to confine ourselves to sentences of modal degree 1, we shall however not proceed in this way, but rather use the operator K in the axiomatic characterisation of the relation $\leqslant_s$ in Section 6.4. In the semantics for $\leqslant_s$ we do not need K but only f.[9]

Let $\mathfrak{P}_3$ be the language obtained from $\mathfrak{A}$ by stipulating that $A \leqslant_s B$ and $K(A, B)$ be sentences of $\mathfrak{P}_3$ if A and B are sentences of $\mathfrak{A}$. We arrive then at the following concept of an interpretation of $\mathfrak{P}_3$:

(D4.1) An interpretation of $\mathfrak{P}_3$ is a quintuple $\langle I, i_0, f, .\leqslant, \Phi \rangle$ such that

(1) I is a set of worlds with $i_0 \in I$.

(2) f is a function from subsets of I to subsets of I such that

(a) $f(X) \subset X$

(b) $X \subset Y \wedge f(X) \neq \Lambda \supset f(Y) \neq \Lambda$

(c) $X \subset Y \wedge f(Y) \cap X \neq \Lambda \supset f(X) = f(Y) \cap X$

(d) $i_0 \in f(I)$.

(3) $.\leqslant$ is a relation on I as in (D3.3–3) with $S = U_X f(X)$.

(4) Φ_i is a function from sentences of $\mathfrak{P}_3$ into $\{t, f\}$ so that

(a) $\Phi_i(\neg A) = t$ iff $\Phi_i(A) = f$ and $\Phi_i(A \wedge B) = t$ iff $\Phi_i(A) = \Phi_i(B) = t$

(b) $\Phi_i(A \leqslant_s B) = t$ iff $\bigwedge i(i \in f(A) \supset \bigvee k(k \in f(B) \wedge i. \leqslant k))$
(c) $\Phi_i(K(A, B)) = t$ iff $f(B) \subset [A]$.

While the conditions (2a)–(2c) for g in (D3.20) derive from our definition $g(X) = \{i : i \in X \cap S \wedge \bigwedge j(j \in X \supset j. \leqslant i)\}$ and the properties of $. \leqslant$ according to (D3.3), the corresponding conditions (2a) to (2d) in (D4.1) are based on our intuitive interpretation in case of f. (2a) and (2b) are trivial, (2c) postulates coherence in the choice of normal worlds. It is equivalent to $f(Y) \cap X \neq \Lambda \supset f(X \cap Y) = f(Y) \cap X$, i.e. if X is possible on condition that Y, then the normal $X \cap Y$-worlds are the normal Y-worlds in which X holds. (2d) is necessary for the principles $LA \supset A$ and $K(A, B) \supset (B \supset A)$ to hold. We may define factual necessity L by

(D4.2) $LA := K(A, \neg A)$,

since then

(4.3) $\Phi_i(LA) = t$ iff $R \subset [A]$, where $R := U_X f(X)$.

NA, normative necessity, may be defined in analogy to (D3.5) by $NA := \neg A \leqslant_s \bot$.

We do not postulate any close connections between f and g, as for instance $\bigwedge ij(i \in f(X) \wedge j \in f(X) \supset i. = j)$. Differences in value do not generally correspond to differences in conditional necessity. But our postulate $R = S$, which implies $NA \equiv LA$ so that the (absolutely) impossible worlds are the normatively inevaluable worlds, seems to be harmless. It could, of course, easily be dropped, but simplifies the formalism somewhat.

According to this analysis $A \leqslant_s B$ is a notion of intrinsic betterness that is based on excluding farfetched possibilities. A statement $B <_s A$ is not based on the fact that there is an A-world, i.e. a possibility of A coming about – no matter how remote it may be – that is better than all B-worlds, but on the fact that among the normal ways, in which A may come about, there is one that is better than all the normal ways, in which B may come about. For many applications this is a more realistic concept of preference than intrinsic betterness. But it is not a *better* notion of preference but a *different* notion of preference, since it relies on assumptions on what normally happens. What is compared with $\leqslant_s$ is not the value of A and B as such but the value of what normally comes of A and B.

The definition of the 2-place relation $\leqslant_s$ carries over to that of the 4-place relation as in the last section. And the classificatory and metric concepts G_s and V_s stand in the same relation to $\leqslant_s$ as G_0 and V_0 to $\leqslant_0$. Therefore we need not go into that but can turn to the relation $\leqslant_p$ directly.

5. *Expected Value*

In comparing two propositions X and Y we may wish not only to take into account the best (or worst) worlds in X and Y but all the X- and Y-worlds, and say, that X is better than Y if *most* X-worlds are better than *most* Y-worlds, or if the *average* X-world is better than the *average* Y-world.

Let us assume for the moment that the set I of worlds is finite. In order to speak about averages we have to introduce real numbers $v(i)$ for the values of the worlds $i \in I$. This may be done, as indicated in Section 3, by a metrisation of the relation $i, j. \leqslant k, l$ as used in (D3.10). Then the value $V(X)$ can be determined as the weighted mean of the $v(i)$ with $i \in X$:

$$(5.1) \qquad V(X) = \frac{\sum_{i \in X} v(i) \cdot m(i)}{\sum_{i \in X} m(i)}.$$

The weights $m(i)$ are used so that $V(X)$ may depend on the values $v(i)$ for some $i \in X$ more than for others. As indicated in Section 3 we may also define a notion of intrinsic value this way. But if we take the weights $m(i)$ to be the probabilities $p(i)$ of the worlds i, so that $V(X)$ depends on the values of the probable X-worlds more than on those of the improbable ones, then we arrive at the notion $\leqslant_p$ as sketched in Section 2. p has to be a subjective probability here, but in case of normative values p will not refer to the beliefs of a single individual but to common beliefs in the population which accepts the norms.[10]

If v is defined on I and a probability measure p on the power set of I we obtain instead of (5.1)

$$(5.2) \qquad V_p(X) = \frac{1}{p(X)} \sum_{i \in X} v(i) \cdot p(i) \quad \text{for} \quad p(X) > 0.$$[11]

This is equivalent to $V_p(X) = \sum_{i \in I} v(i) \cdot p(i, X)$ where $p(i, X)$ is the probability of i given that X, i.e. $p(i, X)=p(i)/p(X)$ for $i\in X$ and $p(X)>0$. $V_p(X)$, therefore, is the expectation of v on condition that X, and in this sense we call $V_p(X)$ the *expected value* of X.

From (5.2) we obtain the fundamental law

(5.3) $$V_p(X+Y)=\frac{V_p(X)\cdot p(X)+V_p(Y)\cdot p(Y)}{p(X+Y)}$$ for

$p(X+Y)>0$, where $X+Y$ represents the union of disjoint sets X and Y.

$V_p(X)$ is not defined for $p(X)=0$ by (5.2) but we shall stipulate that V_p is also determined in these cases and that for $p(X)=p(Y)=0$ we have $V_p(X)=V_p(Y)$. The value of V_p for such practically impossible propositions can be left open.

We may also define conditional expected value by

(D5.4) $$V_p(X, Y)=\frac{1}{p(X, Y)}\sum_{i\in X} v(i)\cdot p(i, Y) \quad \text{for} \quad p(X, Y)>0.$$

Then we obtain

(5.5) $$V_p(X, Y)=V_p(X\cap Y).$$

From the metric concept of expected value defined on the basis of metric concepts v and p of intrinsic value and probability of worlds, we obtain comparative concepts of expected value by

(D5.6) (a) $A, B\leqslant_p C, D := V_p(A)-V_p(B)\leqslant V_p(C)-V_p(D)$

(b) $A\leqslant_p B := V_p(A)\leqslant V_p(B)$,

where $V_p(A)=V_p([A])$.

If we introduce a functor for probabilistic necessity by

(5.7) $$\Phi_i(LA)=t \quad \text{iff} \quad p([A])=1,$$

we obtain from (5.3) the following principles

(5.8) (a) $A=_p B\supset A+B=_p A$

(b) $L\neg A\supset A+B=_p B$

(c) $A<_p B\wedge\neg L\neg B\supset A<_p A+B$

(d) $A <_p B \wedge \neg L \neg A \supset A + B <_p B$
(e) $A =_p B \wedge \neg (C =_p A) \wedge \neg L \neg C \wedge A + C =_p B + C \supset$
$\supset A + D =_p B + D$.[12]

(c) and (d) express what may be called the *strong mean-value-principle*. This principle distinguishes the relation $\leqslant_p$ from $\leqslant_0$. For both of them only the weak *mean-value-principle* $A \leqslant . B \supset A \leqslant . A \vee B \leqslant B$ holds; in fact we have $(A \vee B = . A) \vee (A \vee B = . B)$ for the latter two relations.

As we have pointed out in Section 2 a metric concept should be introduced by way of metrisation of a comparative concept. Therefore the question arises, how we can determine the 4-place relation $X, Y \leqslant_p Z, U$ between propositions so that a representation theorem can be proved, that for all such relations there exist functions v and p so that V_p, defined as in (5.2), is unique up to positive transformations, and satisfies (D5.6–a).

Bolker (1966) and (1967) has shown that, if the 2-place relation $\leqslant_p$ satisfies conditions (5.8) and some others and if there is a function V_p obeying (D5.6–b), suitable functions v and p exist.[13] The gap in the metrisation theorem, that the existence of such a function V_p has to be presupposed, may be closed by starting from the 4-place relation $\leqslant_p$ and postulating that it be an algebraic difference system. Then the existence of V_p is guaranteed by the representation theorem in Krantz *et al.* (1971), p. 151.

If a propositional relation $X, Y \leqslant_p Z, U$ is determined in this way, we have a semantics for the operator $\leqslant_p$ and may define $\Phi_i(A, B \leqslant_p C, D) = t$ iff $[A], [B] \leqslant_p [C], [D]$. But the conditions for $X, Y \leqslant_p Z, U$ are not all representable in our propositional languages so that propositional logics for $\leqslant_p$ will be incomplete.

We define the classificatory concepts of goodness G_p again by

(D5.9) (a) $G_p A := \neg A <_p A$
(b) $G_p(A, B) := \neg A \wedge B <_p A \wedge B$.

This weak goodness does not coincide with strong goodness anymore. If we let $G^*_p(A)$ be true iff $p([A]) > 0$ and $\bigwedge X(p(X) > 0 \supset X \leqslant_p [A])$, then, in case there are optimal worlds and for the set Z of them we have $p(Z) > 0$, we have $p([A] \cap Z) = p([A])$, i.e. $[A]$ contains almost only best worlds. We cannot now define OA by $G^*_p A$, but only by $\bigwedge X(G^*_p(X) \supset$

$\supset X \subset [A]$). This would then be essentially (i.e. taking no account of the practically impossible worlds) the same concept as that defined in (D3.17).

This concludes our semantic analysis of the three basic types of normative concepts. In the last section we formulate logics for the $_0$ – and $-_s$ terms. For the reasons just indicated we do not formulate logics for the p-terms.

6. *Logics*

The following logical systems are sound and complete. The method of completeness proofs, which are very simple, is illustrated for the case of *P*3. Completeness proofs for *P*1, *P*2 and for conditional logic (*P*4.1–*P*4.6 and *P*4.13) may be found in Lewis (1973).[14]

6.1. *Dyadic deontic logic.* Let *PO* be a propositional calculus. Then *P*1 is to be *PO* plus the following axioms and rule:

(1) $O(A, A)$
(2) $NA \supset O(A, B)$
(3) $N(A \supset B) \wedge O(A, C) \supset O(B, C)$
(4) $O(A, B) \wedge O(C, B) \supset O(A \wedge C, B)$
(5) $\neg O(\neg B, A) \supset (O(C, A \wedge B) \equiv O(B \supset C, A))$
(6) $NA \supset A$
(7) $A \vdash NA$

Here N is defined by $NA := O(A, \neg A)$. *P*1, as the following systems, contains the basic system of modal logic (Fey's T and von Wright's M). If we set $OA := O(A, T)$, we obtain the standard system of unconditional deontic logic from *P*1.

In view of (D3.17) *P*1 can also be used as a logic for $G_0(A, B)$ if we substitute $G_0(A, B) \vee N\neg B$ for $O(A, B)$.

6.2. *Dyadic intrinsic preference.* Let *P*2 be *PO* plus the axioms and rule:

(1) $(A \leqslant_0 B) \vee (B \leqslant_0 A)$
(2) $(A \leqslant_0 B) \wedge (B \leqslant_0 C) \supset (A \leqslant_0 C)$
(3) $(A \leqslant_0 B) \supset (A \vee B =_0 B)$

(4) $NA \supset A$
(5) $N(A \supset B) \supset (A \leqslant_0 B)$
(6) $A \vdash NA$.

Here we use the definition (D3.5): $NA := \neg A \leqslant_0 \bot$.

6.3. *Relative intrinsic preference.* $P3$ is to be PO plus

(1) $(A, B \leqslant_0 C, D) \vee (C, D \leqslant_0 A, B)$
(2) $(A, B \leqslant_0 C, D) \wedge (C, D \leqslant_0 E, F) \supset (A, B \leqslant_0 E, F)$
(3) $(A, B \leqslant_0 C, D) \supset (A, C \leqslant_0 B, D)$
(4) $(A, B \leqslant_0 C, D) \supset (D, C \leqslant_0 B, A)$
(5) $A \leqslant_0 B \supset A \vee B =_0 B$
(6) $NA \supset A$
(7) $N(A \supset B) \supset A \leqslant_0 B$
(8) $A \vdash NA$.

Here we use the definition $NA := \neg A \leqslant_0 \bot$ again and $A \leqslant_0 B :=$ $:= A, B \leqslant_0 A, A$.

$P3$ differs from $P2$ only in that the characteristic axioms for dyadic weak orderings ($P2.1$ and $P2.2$) are replaced by those ($P3.1$–$P3.4$) for relative weak orderings.

6.4. *Normal preference.* $P4$ is to be PO plus

(1) $K(A, A)$
(2) $NA \supset K(A, B)$
(3) $N(A \supset B) \wedge K(A, C) \supset K(B, C)$
(4) $K(A, B) \wedge K(C, B) \supset K(A \wedge C, B)$
(5) $\neg K(\neg B, A) \supset (K(C, A \wedge B) \equiv K(B \supset C, A))$
(6) $K(A, B) \supset (B \supset A)$
(7) $(A \leqslant_s B) \vee (B \leqslant_s A)$
(8) $(A \leqslant_s B) \wedge (B \leqslant_s C) \supset (A \leqslant_s C)$
(9) $(A \leqslant_s B) \wedge \neg K(\neg B, A \vee B) \supset (A \vee B =_s B)$
(10) $K(A, B) \wedge K(B, A) \supset A =_s B$
(11) $N \neg A \supset A \leqslant_s B$
(12) $A \leqslant_s \bot \supset N \neg A$
(13) $A \vdash NA$.

Here NA, which in view of ($P4.11$) and ($P4.12$) is equivalent to $\neg A \leqslant_s \perp$, may be defined according to (D4.2) and (D3.5) by $K(A, \neg A)$.

6.5. *Normal relative preference.* $P5$ is obtained from $P4$, as $P3$ from $P2$, by replacing the characteristic axioms for dyadic weak orderings by those for relative weak orderings. $P5$, therefore, is to be $P4$ minus $P4.7$ and $P4.8$, plus $P3.1$–$P3.4$.

6.6. *Completeness of* $P3$. If A is not provable in $P3$ and $\Gamma(A)$ is the set of propositional constants occurring in A, there is a set $\mathfrak{B}$ of sentences of $\mathfrak{P}_2$ with $\neg A \in \mathfrak{B}$ that is $P3$-consistent and $\Gamma(A)$-maximal, i.e. from $\mathfrak{B}$ no contradiction can be derived in $P3$ and if a sentence B with $\Gamma(B) \subset \Gamma(A)$ is not in $\mathfrak{B}$, then $\mathfrak{B} \cup \{B\}$ is $P3$-inconsistent. This is proved as usual by Henkin's procedure. Let I be a set of indices for all interpretations of the basic language $\mathfrak{A}$ satisfying (D3.10–4a), and $E(i)$ the sentences of the form $(\neg)\, p_1 \wedge \ldots \wedge (\neg)\, p_n$ (with $\Gamma(A) = \{p_1, \ldots, p_n\}$) satisfied by Φ_i. If there should be no $E(i)$ with $(E(i) \leqslant_0 \perp) \in \mathfrak{B}$, we add a new constant q to $\Gamma(A)$ and construct $\mathfrak{B}$ so that it also contains $(E(j) \wedge q) \leqslant_0 \perp$ for some j. We may presuppose, therefore, in the following that $\mathfrak{B}$ and $\Gamma(A)$ are so determined that there is an i with $(E(i) \leqslant_0 \perp) \in \mathfrak{B}$.

Let i_0 be that i for which $\Phi_i(q) = t$ iff $q \in \mathfrak{B}$ for all constants q. We define $S := \{i : (\perp <_0 E(i)) \in \mathfrak{B}\}$. We have constructed $\mathfrak{B}$ so that $\bar{S} \neq \Lambda$ and determined i_0 so that $i_0 \in S$. Otherwise we would obtain from $(E(i_0) \leqslant_0 \perp) \in \mathfrak{B}$ $N \neg E(i_0) \in \mathfrak{B}$, and hence with $P3.6$ $\neg E(i_0) \in \mathfrak{B}$, in contradiction to our choice of i_0.

We further define $i, j . \leqslant k, l := (E(i), E(j) \leqslant_0 E(k), E(l)) \in \mathfrak{B}$. Then this relation $. \leqslant$ satisfies the conditions (3a)–(3d) of (D3.10) on the basis of $P3.1$–$P3.4$. It also satisfies (3e): If $(E(i) \leqslant_0 \perp) \in \mathfrak{B}$, then $(E(i) \leqslant_0 E(j)) \in \mathfrak{B}$, for from $N \neg E(i)$ we obtain $N(E(i) \supset E(j))$ and by $P3.7$ $E(i) \leqslant_0 E(j)$. If $(E(i) \leqslant_0 \perp) \in \mathfrak{B}$ and $(\perp <_0 E(j)) \in \mathfrak{B}$ then, in view of $P3.1$ and $P3.2$, we have $(E(i) <_0 E(j)) \in \mathfrak{B}$; therefore (3$f$) holds. And (3$g$) is true, since there is always a maximal $i \in X$ if $X \neq \Lambda$, because there are only finitely many differences between the worlds in I relative to $. \leqslant$.

We have used two theorems of $P3$ tacitly here:

(T3.1) (a) $A \equiv A' \vdash (A \leqslant_0 B) \equiv (A' \leqslant_0 B)$

(b) $B \equiv B' \vdash (A \leqslant_0 B) \equiv (A \leqslant_0 B')$

Proof. (a) From $A \supset A'$ we obtain $N(A \supset A')$ by $P3.8$, hence $A \leqslant_0 A'$ by $P3.7$, and from $A' \leqslant_0 B$ then $A \leqslant_0 B$ by $P3.2$. From $A' \supset A$ we obtain $A' \leqslant_0 A$ and $A' \leqslant_0 B$ from $A \leqslant_0 B$. (b) is proved in the same way.

(T3.2) $N(A \supset B) \wedge NA \supset NB$

Proof. From NA wo obtain $\neg A \leqslant_0 \perp$, from $N(A \supset B)$ $\neg(A \supset B) \leqslant_0 \perp$, and with (T3.1)$\neg(\neg B \supset \neg A) \leqslant_0 \perp$, hence $N(\neg B \supset \neg A)$. With $P3.7$ we then have $\neg B \leqslant_0 \neg A$ and therefore $\neg B \leqslant_0 \perp$, i.e. NB.

Below we also use the theorem

(T3.3) (a) $(A =_0 A') \supset ((A, B \leqslant_0 C, D) \equiv (A', B \leqslant_0 C, D))$
(b) $(B =_0 B') \supset ((A, B \leqslant_0 C, D) \equiv (A, B' \leqslant_0 C, D))$
(c) $(C =_0 C') \supset ((A, B \leqslant_0 C, D) \equiv (A, B \leqslant_0 C', D))$
(d) $(D =_0 D') \supset ((A, B \leqslant_0 C, D) \equiv (A, B \leqslant_0 C, D'))$.

Proof. (a) From $A, B \leqslant_0 A, B$ ($P3.1$, $P3.2$) we obtain $A, A =_0 B, B$ by $P3.3$. $A =_0 A'$ therefore implies $A, B =_0 A', B$ ($P3.3$), hence $(A, B \leqslant_0 C, D) \equiv (A', B \leqslant_0 C, D)$. (b) is obtained the same way, and (c), (d) from (a), (b) by $P3.4$.

If we define the Φ_i also for sentences of the form $A, B \leqslant_0 C, D$ according to D3.10–4b, then $\mathfrak{J} = \langle I, i_0, S, .\leqslant, \Phi \rangle$ is an interpretation of $\mathfrak{P}_2$ in the sense of that definition. $\mathfrak{J}$ satisfies exactly these sentences B with $\Gamma(B) \subset \subset \Gamma(A)$ that are elements of $\mathfrak{B}$, i.e. this is true for Φ_{i_0}. This is trivial for all sentences of the basic language $\mathfrak{A}$.

If $\Phi_{i_0}(A, B \leqslant_0 C, D) = t$ there exist i, j, k, l with $i \in g'(A)$, $j \in g'(B)$, $k \in g'(C)$, $l \in g'(D)$ and $i, j. \leqslant k, l$, hence $(E(i), E(j) \leqslant_0 E(k), E(l)) \in \mathfrak{B}$. From this we can infer $(A, B \leqslant_0 C, D) \in \mathfrak{B}$. For if $E(i_1) \vee \ldots \vee E(i_m)$ is the full disjunctive normal form of A, we obtain from D3.11–b $i_r . \leqslant i$, and hence $(E(i_r) \leqslant_0 E(i)) \in \mathfrak{B}$ for all $r = 1, \ldots, m$. If $(E(i) \leqslant_0 \perp) \in \mathfrak{B}$, we have then $(A =_0 E(i)) \in \mathfrak{B}$, in view of $P3.5$ and $P3.7$, and this is also true in case of $(\perp <_0 E(i)) \in \mathfrak{A}$ in view of $P3.5$, for then $i \in [A]$. Since we also have $B =_0 E(j)$, $C =_0 E(k)$, and $D =_0 E(l)$, we obtain $(A, B \leqslant_0 C, D) \in \mathfrak{B}$ by T3.3.

In the same manner it follows from $\Phi_{i_0}(A, B \leqslant_0 C, D) = f$ that $A, B \leqslant_0 C, D$ is not in $\mathfrak{B}$.

Φ_{i_0} therefore satisfies all sentences in $\mathfrak{B}$, and therefore $\neg A$, so that the sentence A, which was not provable in $P3$, is not valid.

The same method illustrated here for $P3$ may also be used for completeness-proofs for the other logical systems.

Regensburg

NOTES

[1] Cf. Kutschera (1973), 1.5.
[2] To give an example from subjective preference: If I say that I prefer to buy a used car (A) on condition that I am short of money (B) to buying it on condition that I have enough then this is to be represented by $A \wedge \neg B, \neg B <. A \wedge B, B$, not by $A, \neg B <. A, B$ or by $A \wedge \neg B <. A \wedge B$. I may consider buying new cars to be preferable to buying used ones in principle ($A <. \neg A$) and only to be advisable in case one is short of money, though I certainly consider the difference in value between A and $\neg A$ as negligible relative to the difference of B and $\neg B$. Therefore we have $A \wedge B <. A \wedge \neg B$. $A, \neg B <. A, B$ may be true, but is not appropriate, since we do not compare A as such, but only $A \wedge B$, relatively to B, with $A \wedge \neg B$ (not A as such) relatively to $\neg B$.
[3] A relative concept of goodness would be goodness as to some standard of comparison.
[4] For the sake of brevity we use the logical symbols $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$ of our object language besides the quantifiers $\bigwedge$ and $\bigvee$ also as metalinguistic symbols.
[5] For some other definitions cf. Danielsson (1968) and Kutschera (1974).
[6] Cf. also Lewis (1973), p. 121 for the role of the limit-assumption, and the completeness proof in Section 6.6 below.
[7] For the equivalence of the $.\leqslant$-semantics with the g-semantics under the limit-assumption cf. Lewis (1973), 2.7.
[8] This interpretation is discussed in Kutschera (1975).
[9] For the problems connected with the functor F cf. Lewis (1973), 2.8.
[10] Such common beliefs play a role in law, e.g. if a judge maintains that the defendant in a situation S should have reckoned with a certain contingency X, since it is common knowledge that in situations of the type S it has to be expected that X happens.
[11] We write $p(i)$ for $p(\{i\})$. In case of infinitely many worlds integrals have to be used instead of sums.
[12] $A+B$ is used instead of $A \vee B$ to indicate that $[A]$ and $[B]$ are disjoint.
[13] Cf. also Jeffrey (1965).
[14] For $P4$ cf. his system VW. The proofs he gives are much simplified for our language without iterated applications of modal operators.

BIBLIOGRAPHY

Åqvist, L., 1973, 'Modal Logic with Subjunctive Conditionals and Dispositional Predicates', *Journal of Philosophical Logic* **2** (1973), 1–76.

Bolker, E., 1966, 'Functions Resembling Quotients of Measures', *Transactions of the American Mathematical Society* **124** (1966), 292–312.

Bolker, E., 1967, 'A Simultaneous Axiomatisation of Utility and Subjective Probability', *Philosophy of Science* **34** (1967), 333–340.

Danielsson, S., 1968, *Preference and Obligation*, Uppsala 1968.

Jeffrey, R. C., 1965, *The Logic of Decision and Action*, New York 1973.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A., 1971, *Foundations of Measurement*, New York 1971.

Kutschera, F. v., 1973, *Einführung in die Logik der Normen, Werte und Entscheidungen*, Freiburg i.B. 1973.

Kutschera, F. v., 1974, 'Normative Präferenzen und bedingte Obligationen', to appear in H. Lenk (ed.), *Normenlogik*, München, 1974, pp. 137–165.

Kutschera, F. v., 1975, 'Indicative Conditionals', to appear in *Theoretical Linguistics*. **1** (1974).
Lewis, D., 1973, *Counterfactuals*, Oxford 1973.
Stalnaker, R. C., 1968, 'A Theory of Conditionals', in N. Rescher (ed.), *Studies in Logical Theory*, Oxford 1968, pp. 98–112.