A SEMANTICS FOR A LOGIC OF 'BETTER'

M. J. Cresswell

In [1] we defined a deontic operator **O** in terms of the truth or falsity of its argument in good or bad worlds in a model. (1) Now reasons such as those found in [2] lend weight to the view that the tripartite classification of situation into good, bad and indifferent is too narrow. Surely, one might say, propositions can be of varying worth.

The most natural way of tackling this formally seems to be by a logic having some kind of betterness operator. Several such systems have been developed (2) and have been characterized in one way or another. We propose to put forward a definition of validity based on models with sets of worlds and to have a look at some of these systems in the light of this. We shall also make one or two remarks about possible axiomatizations for such a system.

One difficulty in evaluating propositions is that a proposition can be true in many different worlds, some good and some bad. And the value of truth functions seems to be independent of the value of their arguments. E.g. take the propositions 'x robs a bank' and 'x endows a bishopric' and their conjunction 'x robs a bank and endows a bishopric'. Here if we can work out the value of the conjunction at all we would probably incline to say it was somewhere between the value of its members. But take the propositions (3) 'John loves Mary' and 'Mary loves John'. Although the two constituents are

(1) We assume some familiarity with the terminology of [1], in particular the concept of a model as a set W of worlds and of an assignment V within such a model. In what follows the truth functions (and necessity where applicable) are evaluated in the usual way. We also continue to assume that deontic operators have propositional arguments.

(2) E.g. [3], [4] and [5].

(3) I am indebted to Professor G. E. Hughes for this example and for discussion on the problems it raises but he is not to be blamed for any conclusions I have come to.

very difficult to evaluate one is inclined to say that their conjunction 'John loves Mary and Mary loves John' is better than either of them. In so far as we can make sense of this at all it could be that the reason is that a world in which both Mary loves John and John loves Mary will almost certainly be better than a world in which either is true on its own. (4) I.e. if we take all the worlds in which John loves Mary this set will contain many bad ones (the ones in which Mary does not love John) though of course it will contain some good ones, and if we take all the worlds in which Mary loves John this will contain many bad ones (the ones in which John does not love Mary). If now we take the worlds in which both are true this set will (in respect of John and Mary's loving) be better since it won't contain the (bad) worlds in which the love is unrequited. We use this kind of model where x_1 is a good world and x_2, x_3 are bad worlds and x_4 is indifferent,

x_1	x_2	x_3	x_4
p	p	$\sim p$	$\sim p$
q	$\sim q$	q	$\sim q$

In this model p is indifferent (for it is true in one good world, and one bad world). Similarly q is indifferent (for it is true in one good and one bad world) but $(p \,.\, q)$ is only true in x_1, i.e. is only true in a good world, and so is better than either p or q. On the other hand $(p \,.\, \sim q)$ is worse than either of them for it is only true in a bad world (and similarly $(q \,.\, \sim p)$).

If we were to give worlds (positive and negative) values then we could determine the value of a proposition by determining the total value of the worlds in which it is true (clearly this could only apply to a finite W). Let us assume each $x_i \in W$ is assigned a positive (for good) or negative (for ill) integral

(4) In his discussion of his axiom A7 ([3] p. 29) Halldén seems to be giving an argument of this sort for its validity. Perhaps the real moral of all this is that it is only total situations and not propositions which can sensibly be compared but be that as it may it can still, we think, be pointful to consider a semantics in which propositions can be compared.

value. (integers including 0) Where $P(x_i)$ is the value of x_i and $U \subseteq W$ then $P(U)$ is the value of the sum of all the $P(x_i)$'s for each $x_i \in U$. This will be a positive or negative integer.

<VWP> is a *B model* iff V is an assignment satisfying V . 1 — V . 4 of [1] p. 179 and,

VB1: for wffs α and β and $x_i \in W$, $V(B\alpha\beta x_i) = 1$ iff, where U is the set of $x_j \in W$ such that $V(\alpha x_j) = 1$ and U' the set of $x_j \in W$, such that $V(\beta x_j) = 1$ then $P(U) > P(U')$, otherwise 0.

A formula is *B-valid* iff for every B-model <VWP> $V(\alpha x_i) = 1$ for every $x_i \in W$.

The assignment of integral values to worlds seems like an advance on the crude 'good' and 'bad' models but interestingly enough it proves formally unnecessary.

Let <VWG> be a *B' model* iff W is a finite set of worlds, $G \subseteq W$ and V an assignment satisfying V . 1 — V . 4 and,

VB2: for wffs α, β and $x_i \in W$, $V(B\alpha\beta x_i) = 1$ iff the number of worlds in G less the number of worlds in W — G in which α is true is greater than the number of worlds in G less the number of worlds in W — G in which β is true.

Given a B model <VWP> we construct a B' model <V'W'G> such that for any wff α, if $V(\alpha x_j) = 0$ for some $x_i \in W$ then $V'(\alpha x_i) = 0$ for some $x_i \in W'$. Since any B' model already is a B model (in which the good worlds have the value + 1 and the bad worlds — 1) this will shew that both kinds of model determine the same set of valid formulae.

Given <VWP> let W' be defined as follows.

Where W is $\{x_1, \ldots, x_n\}$ then where the value of x_i $(1 \leqslant i \leqslant n)$ is k (i.e. $P(x_i) = k$) then if k is positive let $x^1_i, \ldots, x^k_i$ be good worlds in W'. If k is negative let them be bad worlds. Thus $W' = \{x^1_1, \ldots, x^h_h\}$. Let $V'(px^j_i) = V(px_i)$. We shew that $V'(\alpha x^j_i) = V(\alpha x_i)$. Proof by induction on the construction of α. By the definition of V' it holds for the variables. Obviously it holds for truth functions. For modality if $V(L\alpha x_i) = 1$ then $V(\alpha x_j) = 1$ for every $x_j \in W$, hence (induction hypothesis) $V'(\alpha x^k_j) = 1$ for every $x^k_j \in W'$, and $V'(L\alpha x^h_i) = 1$. If $V(L\alpha x_i) = 0$ then for some $x_j \in W$, $V(\alpha x_j) = 0$, $V'(\alpha x^k_j) = 0$, $V'(L\alpha x^h_i) = 0$.

For B, suppose $V(B\alpha\beta x_i) = 1$ then where the sum of the worlds in which α is true is k and in which β is true is h, $k > h$. Now for each world of value l in <VWP> there will be l good worlds if l is positive and l bad worlds if l is negative. The result of adding the values will be the same as adding the good worlds and subtracting the bad ones. Thus $V'(B\alpha\beta x_i^j)$

$= 0$. By a similar argument we have if $V(B\alpha\beta x_i) = 0$ then $V'(B\alpha\beta x_i^j) = 0$ and the induction is proved.

This semantics entails that if a formula $B\alpha\beta$ is true in a model then it is true in every world in the model. It is possible (by analogy with [1] pp. 187-188) to change the modelling to fit the view that different things can be better in different worlds. For we often want to say that something is *better in the circumstances* than something else. Thus we restrict ourselves to counting up worlds which are relevant to the given world (cf. [1] p. 187). This relevance is expressed by a relation betwen worlds xRy could be interpreted as something like 'x is a world value-relevant to y'. We shall not try to analyse the meaning of R but given it we can define **VB3**: $V(B\alpha\beta x_i) = 1$ iff the number of worlds in G *related by R* to x_i in which α is true less the number of worlds in $W - G$ related to x_i in which α is true is greater than the number of worlds in G related to x_i in which β is true less the number of worlds in $W - G$ relevant to x_i in which β is true.

This makes the semantics more complicated and for the present we shall restrict ourselves to semantics without a relation R.

We shall also restrict ourselves to the purely deontic system. We first note that where γ results from α (or β) by replacing some $B\delta\eta$ in α (or β) by $(p \supset p)$ and γ' results by replacing it by $\sim(p \supset p)$ then

$B\alpha\beta \equiv ((B\delta\eta \,.\, B\gamma\beta) \vee (\sim B\delta\eta \,.\, B\gamma'\beta)$

$B\alpha\beta \equiv ((B\delta\eta \,.\, B\alpha\gamma) \vee (\sim B\delta\eta \,.\, B\alpha\gamma')$ are both valid. This is because if a deontic formula is true it is so in every world and hence has the same truth-value (in every model) as $(p \supset p)$ and if false as $\sim(p \supset p)$. These equivalences enable us to bring embedded deontic formulae outside the scope of other deontic operators and reduce the formula to a first-degree one, *viz* a truth function of PC formulae and formulae of the form $B\alpha\beta$ where α and β are PC formulae. (5)

This means that we can without loss of generality restrict ourselves to the kind of systems considered in ([3], [4] and [5]). Of these one which gives a semantics is Åqvist's. The theorems of Åqvist's BD are precisely those which are true in all two world B' models where both worlds are good. For we have the following correspondence between an assignment V' in Åqvist's semantics and an assignment V in a B' model where $W = G = \{x_1, x_2\}$

$$V'(p) = 1 \text{ iff } V(px_1) = 1 \text{ and } V(px_2) = 1$$
$$V'(p) = 1/2 \text{ iff } V(px_1) = 1 \text{ and } V(px_2) = 0 \text{ or } V(px_1) = 0 \text{ and } V(px_2) = 1$$
$$V'(p) = 0 \text{ iff } V(px_1) = 0 \text{ and } V(px_2) = 0.$$

This is why the value $K^1/_2{}^1/_2$ cannot be determined ([4] p. 91) since if p and q are each true in exactly one world, it might be the same one (in which case $p \,.\, q$ is true in exactly one and so $= 1/2$) or different ones (in which case $p \,.\, q$ is true in neither and so $= 0$). Åqvist's matrix for B also satisfies this. Other finite, indeed deontic systems (e.g. [6], [7]) can no doubt be analysed by similar restrictions on the model.

Moving to those which do not seem to use a finite matrix we find that of the principles ([5] p. 40) on which von Wright bases his preference logic the first two are clearly valid (from the properties of ordering) and so is the third *viz* $Bpq \equiv B(p \,.\, \sim q)(q \,.\, \sim p)$ for suppose that in any B' model the number of good worlds in which p and q are both true (i.e. the number of worlds $(p \,.\, q)$ is true) is n and the number of bad worlds n', suppose that the number of good worlds in which p is true but q false is m and the number of bad worlds m', and suppose that the number of good worlds in which p is false but q true, k and the number of bad worlds k'. I.e. we have the following:

(5) With modality (S5) this reduction procedure amounts to that given in [1] p. 184.

	G	B
$(p \,.\, q)$	n	n'
$(p \,.\, \sim q)$	m	m'
$(q \,.\, \sim p)$	k	k'

Where the value of

$$(p \,.\, \sim q) = m - m'$$
$$(q \,.\, \sim p) = k - k'$$

Now the worlds in which p is true will be precisely those in which either $(p \,.\, q)$ or $(p \,.\, \sim q)$ is true and the worlds in which q is true those in which either $(p \,.\, q)$ or $(q \,.\, \sim p)$ is true, hence the value of $p = (n + m) - (n' + m') = (n - n') + (m - m')$ and the value of $q = (n + k) - (n' + k') = (n - n') + (m - m')$.
Clearly $(n - n') + (m - m') > (n - n') + (k - k')$ iff $(m - m') > (k - k')$

i.e. $Bpq \equiv B(p \,.\, \sim q)\ (q \,.\, \sim p)$

will be true in any B' model.
However of the next two we may falsify the following instances. For (5) $Bpq \supset B(p \,.\, r)\ (q \,.\, r)$.
This is false in the model $W = \{x_1x_2\}$ $G = \{x_1\}$,

$V(px_1) = 1,\ V(qx_1) = 0,\ V(rx_1) = 0$
$V(px_2) = 0,\ V(qx_2) = 1,\ V(rx_2) = 0$

for while p's value is $+1$, q's is -1 (hence Bpq is true) while $(p \,.\, r)$'s is 0 and $(q \,.\, r)$'s is 0 (hence $B(p \,.\, r)\ (q \,.\, r)$ is false.

For (4) $B(p \vee q)\ (r \vee s) \supset B(p \,.\, \sim r \,.\, \sim s)\ (\sim p \,.\, \sim q \,.\, r)$

$W = \{x_1x_2\}$, $G = \{x_1\}$,
$V(px_1) = 1$, $V(qx_1) = 1$, $V(rx_1) = 1$, $V(sx_1) = 1$
$V(px_2) = 0$, $V(qx_2) = 0$, $V(rx_2) = 0$, $V(sx_2) = 1$.

Here the (deontic) value of $p \vee q$ is $+1$ (true in G but not in $W - G$) of $r \vee s$ is 0 (true in G and B) hence $B(p \vee q)\ (r \vee s)$ is true. But the deontic value of $(p \,.\, r \,.\, s)$ is 0 since it is false in both G and W — G and so is $(p \,.\, q \,.\, r)$.

Now von Wright's system is coherent and easily manipulatable. Thus one might suspect that its theorems are these formulae true in a fairly straightforwardly restrictable class of B' models, as Åqvist's was. However this conjecture is made difficult by the failure of rules like the substitution of tautologous equivalents (v. [5] pp. 46-47). As an example we shall shew how the unrestricted use of this rule, when applied to principle (5) reduces the system to triviality. A consequence of principle (5) is

$$Bpq \supset \mathrm{B}(p \,.\, r)\ (q \,.\, r)$$

for r put $(r \,.\, \sim r)$ hence

$$Bpq \supset \mathrm{B}(p \,.\, r \,.\, \sim r)\ (q \,.\, r \,.\, \sim r).$$

But $(p \,.\, r \,.\, \sim r) \equiv (q \,.\, r \,.\, \sim r)$ hence (from $\sim Bpp$ by subs eq) $\sim B(p \,.\, r \,.\, \sim r)\ (q \,.\, r \,.\, \sim r)$ hence $\sim$Bpq. And clearly if $\sim Bpq$ is a theorem of a logic of betterness the system collapses. (This of course does not show that there is anything wrong with von Wright's system but it does show that it could not have a semantics which is at all like ours.)

Halldén's hypothesis of comparability is valid, [3] p. 27 and so all his axioms are (of Theory A, chapter II). On p. 41 he gives some arguments against requiring there to be as many good models as bad. Halldén's theory B axioms are all valid. Further his matrix (p. 71) is simply a B' one where there are as many good worlds as bad. (Halldén proves his matrix is characteristic for theory B.)

We have already discussed reasons why few distributive laws are valid. These reasons rule out all the formulae rejected by Halldén in Chapter III.

The ethical status of $(p \,.\, \sim p)$ and $(p \vee \sim p)$ is interesting for they would be equal in deontic value precisely in those models which contain as many good worlds as bad. (Since this is a property of a model which is independent of the assignment to the variables we get the 'semantic incompleteness' mentioned on p. 71, in the same way that we have it in S2 because the set of normal worlds ([8] p. 275) in an S2 model is likewise independent of the particular assignment.)

Halldén's system A is weak enough for our model theoretic requirements but whether it is strong enough is an open question.

Victoria University of Wellington. M. J. CRESSWELL

REFERENCES

[1] M. J. CRESSWELL, 'Some Further Semantics for Deontic Logic' *Logique et Analyse* no. 38 (1967) pp. 179-191.

[2] J. O. URMSON, 'Saints and Heroes,' *Essays in Moral Philosophy* (ed. A.I. Melden) Seattle 1958, pp. 198-216.

[3] Sören HALLDÉN, *On the Logic of Better* Lund, 1957.

[4] Lennart ÅQVIST, 'A Binary Primitive in Deontic Logic' *Logique et Analyse* no. 19 (1962) pp. 90-97.

[5] G. H. VON WRIGHT, *The Logic of Preference* Edinburgh, 1963.

[6] Lennart ÅQVIST, 'Postulate Sets and Decision Procedures for Some Systems of Deontic Logic' *Theoria* Vol. 29 (1963) pp. 154-175.

[7] — 'Deontic Logic Based on a Logic of "Better"' *Acta Philosophica Fennica* (1963) *Modal and Many-valued Logics* pp. 285-289.

[8] G. E. HUGHES and M. J. CRESSWELL, *An Introduction to Modal Logics*, London, 1968.

Errata in [1]

p. 186 ⊥ 26, 27, 31 for $x_{h+1} \in G$ read $x_{n+i} \in G$
p. 190 ⊥ 11 for Γ_j read Γ_i
p. 187 after ⊥. 20 (after line (4) of proof) add
OL3.4 (corollary) (5) $O(\alpha \,.\, \sim\beta_1 \,.\, \dots \,.\, \sim\beta_n) \,.\, O(\alpha \vee \sim\beta_1 \vee \dots \vee \sim\beta_n) \supset O\alpha$