1-1-1978

The problem of conditional obligation.

Judith Wagner Decew
University of Massachusetts Amherst

THE PROBLEM OF CONDITIONAL OBLIGATION

A Dissertation Presented

by

JUDITH WAGNER DECEW

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

July 1978

Philosophy

THE PROBLEM OF CONDITIONAL OBLIGATION

A Dissertation Presented

By

JUDITH WAGNER DECEW

Approved as to style and content by:

Fred Feldman, Chaiman of Committee

Edmund L. Gettier, III, Member

John Robison, Member

Emmon W. Bach, Member

Robert C. Sleigh, Jr., Department Head
Philosophy

July, 1978

ACKNOWLEDGEMENT

This dissertation was written during my tenure as a 1977-78 American Association of University Women Fellow. I gratefully acknowledge their support.

I also sincerely thank the members of my dissertation committee for their patience and help. I am extremely grateful to Edmund Gettier for discussions about many sections of the dissertation, particularly Chapter V. His advice, insights, and encouragement, offered willingly and often at a moment's notice, were invaluable. And I am especially thankful for the continuous help I have recieved throughout the past two years from my director, Fred Feldman. His examples, meticulous comments, and constant encouragement, given cheerfully and promptly, influenced the entire dissertation.

Finally, I thank my parents and my husband, Alan, for their irreplaceable support and understanding, and my son, David, for helping me keep my sense of humor.

ABSTRACT

The Problem of Conditional Obligation

(September 1978)

Judith Wagner DeCew, B.A., University of Rochester

M.A., University of Massachusetts

Ph.D., University of Massachusetts

Directed by: Professor Fred Feldman

Various attempts to formalize a concept of conditional obligation within the standard deontic logic have led to paradoxes, indicating that standard versions of deontic logic are inadequate for expressing those obligations.

It has been suggested, most recently by P.S. Greenspan, that the imposition of temporal restrictions on obligation statements may provide solutions to these paradoxes. I argue in Chapter I that Greenspan has not provided an adequate theory of time and obligation nor a solution to the paradoxes she discusses. Furthermore, I construct a more careful theory of time relativized obligation than has previously been presented. My proposal makes it clear why the addition of temporal limits alone to the standard versions of deontic logic can provide a solution to some, but not all, of the problems associated with conditional obligation.

Many of the difficulties associated with conditional obligation in standard versions of deontic logic appear to arise from the use of the material conditional connective in formalizations of sentences expressing those obligations. And in Chapters II and III, I investigate the possibility of employing a stronger conditional connective to provide an acceptable way of expressing conditional obligations. Syntactic and semantic considerations lead me to conclude that a mere change in the conditional connective can not solve the problems confronted. Moreover, my arguments in Chapter III illustrate the inadequacies of semantic analyses for deontic logic, such as Jaakko Hintikka's and Dagfin Føllesdal and Risto Hilpinnen's, based on the notion of deontic perfection or ideality.

In Chapter IV, I present and evaluate David Lewis' much more plausible semantic analysis for deontic logic and theory of conditional obligation. I conclude that Lewis' proposal is ultimately unsuccessful and emphasize that it is, moreover, an analysis of a concept of conditional ought-to-be.

In contrast, my focus is a concept of conditional ought-to-do. In Chapter V, I present and discuss various features of a theory of obligation developed by Richmond Thomason within his indeterminist time model, and I attempt to build an analogous theory of conditional ought-to-do upon that branching time model. Although I recognize the plausibility of Thomason's model for an analysis of conditional obligation, and conclude that it is unlikely that an adequate theory of conditional obligation can be developed without a semantic theory countenancing the concept of an open future, I detail serious difficulties which arise for any theory of conditional ought-to-do built upon Thomason's theory of obligation.

TABLE OF CONTENTS

INTRODUCTION

Standard versions of deontic logic developed by von Wright, Åqvist, and others allow certain symbolizations of sentences about what ought to be the case, what is permitted, and what is forbidden. Unfortunately, however, within the standard or minimal deontic logic[1] it is possible to derive various paradoxes. The most important and interesting of these paradoxes indicate that the standard deontic logic is inadequate for expressing statements of conditional obligation.

As has often been pointed out, the natural candidate for expressing conditional obligations of the form "Given that p, it ought to be the case that q" (where p and q[2] are schematic sentence letters expressing possible states of affairs) in the standard deontic logic leads to difficulties. $O(p \supset q)$, where O is read "it ought to be the case that," a symbolization originally proposed by von Wright,[3] follows if p is forbidden no matter what q is. That is, we can prove $O(\sim p) \supset O(p \supset q)$. This formula alone may be unobjectionable if we read it as asserting that if $\sim p$ is obligatory and compatible with one's other obligations then $\sim p \vee q$ is also. However, if $O(p \supset q)$ is to be a representation of conditional obligation, then the formula tells us that the forbidden conditionally obligates one to <u>anything</u>. For example, assuming that one ought not to lie, then lying conditionally obligates one to blow up the universe or anything else. Surely this is unsatisfactory.

Furthermore, if standard deontic logic is extended to include mixed formulas[4] then $p \supset Oq$ will be a well formed formula. But $p \supset Oq$, an alternative symbolization proposed by A.N. Prior,[5] is similarly inadequate for conditional obligation sentences. For $\sim p \supset (p \supset Oq)$ is valid. Yet if $p \supset Oq$ is to be interpreted as indicating a conditional obligation for q given that p then whatever is not the case conditionally obligates one to anything. Thus the provability of $O(\sim p) \supset O(p \supset q)$ and $\sim p \supset (p \supset Oq)$ shows that the vacuous truth of a material implication with a false antecedent poses a problem for both $O(p \supset q)$ and $p \supset Oq$ as symbolizations for conditional obligations. In each case too much is conditionally obligatory.

In addition, if either $p \supset Oq$ or $O(p \supset q)$ serves as our symbolization of conditional obligation we will be forced to accept certain invalid inferences. For given $p \supset Oq$ then $p \& r \supset Oq$ holds for any r. Similarly, in the standard deontic logic (SDL), since $p \supset q$ implies $p \& r \supset q$, then $O(p \supset q) \supset O(p \& r \supset q)$ follows for any r.[6] And yet it may be that given condition p, q is obligatory, although given condition p & r, q is not obligatory. For example, given that Smith has borrowed \$50 from Jones, then he is obligated to pay Jones \$50. But given that he has borrowed

$50 from Jones and that Jones forgives the loan, he surely is no longer obligated to pay Jones the $50.[7] In general it is a distinguishing feature of conditional obligations, or obligations dependent on conditions, that they may vary as the conditions vary. To deny this is to misunderstand the logic of what I take to be the most interesting concept of conditional obligation.

The inadequacy of expressing conditional obligations using $O(p \supset q)$ or $p \supset Oq$ has been further emphasized by Roderick Chisholm in the following way. He argues that it is reasonable to agree that these four English sentences are intuitively consistent:[8]

(1) It ought to be that a certain man go to the assistance of his neighbors.

(2) It ought to be that if he does go then he tell them he is coming.

(3) If he does not go then he ought not to tell them he is coming.

(4) He does not go.

Yet the natural symbolizations of these sentences lead to a contradiction. For

(1a) Og

(2a) $O(g \supset t)$

(3a) $\sim g \supset O(\sim t)$

(4a) $\sim g$

yield Ot by an application to (1a) and (2a) of a valid detachment rule (axiom (2) of SDL[9]), according to which

Op and O(p ⊃ q) implies Oq. And (3a) and (4a) yield O(∼t) by modus ponens. Hence we can derive contradictory obligations from (1a) - (4a) and this conflicts with another principle of deontic logic that it is not true to say, of any p, both that p ought to occur and ought not to occur.

The two principles appealed to above, the deontic detachment principle and the principle of non-contradictory obligations, are provable in SDL[10] and intuitively acceptable as well.[11] But if we agree to accept these two principles, how can we explain the contradiction above? Apparently the process of interpreting and formalizing sentences (1) - (4) as (1a) - (4a) was incorrect.

A natural move would be to claim that there is a scope ambiguity in (2) and (3). Although the English sentences appear to be formalizable most naturally as (1a) - (4a), one might argue that the second and third statements actually ought to be symbolized analogously. That is, one might claim that both express conditional obligations and despite the fact that in the English sentences the obligation operator governs the entire conditional in (2) and only the consequent in (3), in an adequate formalization the scope of the deontic operator should be the same for both (2) and (3). Thus we might suggest the following symbolization:

(1a) Og

(2a) O(g ⊃ t)

(3b) O(∼g ⊃ ∼t)

(4a) $\sim g$.

Now it appears that (1a) and (2a) still yield Ot by the deontic detachment rule but we can no longer derive $O(\sim t)$ from (3b) and (4a). However the puzzle can not be solved so easily. For the English sentences (1) - (4) are not only consistent but also logically independent.[12] None is a logical consequence of the others. But (3b) is easily deducible from (1a) given the rule easily derivable in SDL according to which $p \supset q$ implies $Op \supset Oq$.[13] The paradoxes of $O(p \supset q)$ plague us again. For no matter what r is, we can derive $O(\sim g \supset r)$ from Og. Clearly this symbolization does not preserve the independence of the English sentences (1) - (4).

For the same reason, symbolizing our sentences as

(1a) Og

(2b) $g \supset Ot$

(3a) $\sim g \supset O(\sim t)$

(4a) $\sim g$

or as

(1a) Og

(2b) $g \supset Ot$

(3b) $O(\sim g \supset \sim t)$

(4a) $\sim g$

will be unacceptable. For in both cases (2b) follows from (4a). Clearly the puzzle can not be solved merely by paying close attention to the scope of the deontic operator. No symbolizations of (2) and (3) by means of $p \supset Oq$ or $O(p \supset q)$ alone will adequately express Chisholm's four sentences.

Sentence (3) expresses a special type of conditional obligation sentence which Chisholm calls a contrary-to-duty imperative. One ought to go to the assistance of his neighbors. But given that he fails to fulfill that obligation he is obliged not to tell them he is coming. A contrary-to-duty imperative tells an agent what he ought to do given that he has neglected his duty.

In general it is surely possible for condition p to make q obligatory, for condition $\sim p$ to make $\sim q$ obligatory, for p to be unconditionally obligatory and yet for $\sim p$ to occur. But we do not have any adequate formalization of such a contrary-to-duty situation in standard versions of deontic logic.[14]

In order to provide an acceptable way of expressing Chisholm-type paradoxes, contrary-to-duty imperatives in particular, and conditional obligations in general, we might either (i) revise or restrict the traditional symbolizations $O(p \supset q)$ and $p \supset Oq$, perhaps by incorporating temporal restrictions, or (ii) augment the standard deontic logic with a new formalization and interpretation for conditional obligation, perhaps a new conditional connective or a dyadic obligation operator, to escape the paradoxes above.

In this dissertation I shall evaluate the adequacy of each of these moves. The first alternative has been suggested by Lennart Åqvist, Wilfrid Sellars, John Robison, Lawrence Powers, and P.S. Greenspan. The second has been proposed by David Lewis, Bas van Fraassen, Bengt Hansson, and Dagfin Føllesdal and Risto Hilpinnen. I shall argue that attempts to pursue (i) alone, to the extent that they avoid or solve the paradoxes, do so at the expense of providing an analysis of a concept different from that concept which I take to be conditional obligation. While I conclude that (ii) is the most promising direction to pursue, I suggest that the current analyses of that type need to be revised to explicate that notion of conditional obligation most crucial for normative ethics.

NOTES TO INTRODUCTION

1. See Appendix I for an explanation of my use of the term "standard deontic logic."

2. I shall adopt the common convention of omitting quotation marks and allowing symbols and formulas to be names of themselves. In each case the meaning is clear from the context.

3. "Deontic Logic," Mind 60 (1951), pp. 1-15.

4. This extension is unproblematic as long as deontic operators are prefixed to sentences rather than names of acts.

5. "The Paradoxes of Derived Obligation," Mind 63 (1954), pp. 64-65.

6. That is, we have $p \supset q$ implies $Op \supset Oq$. This follows in the standard deontic logic (SDL) directly from rule (R) and axiom (2), (see Appendix I). For given $p \supset q$ we have $O(p \supset q)$ by (R) and then $Op \supset Oq$ by (2).

7. See David Lewis' Counterfactuals, Harvard University Press, 1973, pp. 102-103 for another example.

8. "Contary-to-Duty Imperatives and Deontic Logic," Analysis 24 (1963), pp. 33-36. Chisholm notes that he does not believe that this paradox applies to the system set forth by Hector Neri Castañeda. But see P.S. Greenspan's "Practical Reasoning and Deontic Logic: Some Footnotes in Reply to Castañeda," forthcoming in Journal of Philophy.

9. See Appendix I.

10. The first axiom (2) and the second follows from axiom (1) and rules of propositional calculus. That is, $Op \supset \sim O \sim p$ implies $\sim(Op \ \& \ O\sim p)$.

11. But see Bas van Fraassen's "Values and the Heart's Command," Journal of Philosophy 70 (January 11, 1973), pp. 5-19, for an argument that it is our hasty acceptance of the latter which is the root of the difficulty.

12. Cf. Lennart Åqvist, "Good Samaritans, Contrary-to-Duty Imperatives, and Epistemic Obligations," Noûs 1 (December 1967), pp. 365 ff.

13. See note 6. In particular, since we have $g \supset (\sim g \supset \sim t)$, then $Og \supset O(\sim g \supset \sim t)$ follows.

14. Cf. Åqvist's statement of this puzzle in "A Note on Commitment," Philosophical Studies 14 (1963), pp. 22-25.

CHAPTER I

One possible way of dealing with the paradoxes explained above and providing adequate expressions for conditional obligations, especially contrary-to-duty obligations, is to place temporal restrictions on our symbolizations of conditional obligations. Such a time relativized view is motivated by Lawrence Powers' examples in "Some Deontic Logicians" (Noûs, December, 1967). It is also explicitly suggested as a way of avoiding paradoxes of deontic logic by Wilfrid Sellars, Lennart Åqvist, and most recently by P.S. Greenspan in her 1972 Harvard PhD dissertation "Derived Obligation: Some Paradoxes Escaped" and her 1975 Journal of Philosophy paper "Conditional Oughts and Hypothetical Imperatives." The idea is attractive because it seems there can be no contradiction between an absolute obligation to do p at a certain time and an obligation to do $\sim p$ which arises at a later time when a certain condition, perhaps a violation of a prior duty, is fulfilled. Recognition that certain obligations, notably those expressed by contrary-to-duty imperatives, arise only after violation of a duty suggests that a proper characterization of obligations relativized to times might block the derivation of contradictory obligations. This view has indeed been defended but no adequate and complete theory of temporal restrictions on obligations has been formed.

The most recent and explicit attempt to restrict obligations temporally is P.S. Greenspan's. I shall begin by presenting Greenspan's proposal and explaining why she does not provide an acceptable theory of time and obligation nor an adequate escape from the paradoxes. Then building on suggestions by Greenspan and Åqvist I work out a more careful and complete theory of time relativized obligation than has previously been set forth. My proposal makes it clear why the addition of temporal limits on standard symbolizations of conditional obligation statements alone can provide a solution to some but not all of the problems associated with conditional obligations, despite Åqvist's and Greenspan's claims to the contrary.

Greenspan's main project is twofold: (1) to provide solutions to some of the paradoxes arising in the standard system of deontic logic, and (2) to defend some form of a factual detachment rule which allows inferences from $O(p \supset q)$ and p to Oq.[1] Both (1) and (2) can be accomplished, on her view, if proper temporal restrictions are placed on the obligation statements.

Greenspan argues that we are rarely faced with situations in which we ought to perform one particular act. Rather she claims that there are usually several options or possible states of affairs which we may bring about, each of which would be a way of discharging a particular obligation. Thus she claims that the options model or the formal symbolization $O(p \vee q)$, has wide application. Furthermore, since $p \vee q$ is truth functionally equivalent to $\sim p \supset q$ (in her symbols, $\sim p \rightarrow q$), she believes the formulae $O(p \vee q)$ and $O(p \supset q)$ may serve as models for conditional obligation. It is unclear that the wide applicability of $O(p \vee q)$ or $O(\sim p \supset q)$ implies anything whatsoever about its appropriateness for expressions of conditional obligation. Surely the fact that the expression $O(p \supset q)$ contains an obligation operator and a material conditional is not sufficient reason to adopt it as a model for conditional obligation. Nevertheless Greenspan says,

> Indeed, the formula $O(p \rightarrow q)$ and its variants may be stretched to fit any case in which an obligation is conditional upon certain facts--facts about the agent or his situation, acts of other people, or what have you--as in the statement:
>
> (f) I ought to show up in court on the 12th if I am able to do so and the trial is set for that day.
>
> On this interpretation, of course, the condition, or 'if-clause,' is taken as falling within the scope of the deontic operator. . . . in (f), it ought to be the case that, if my trial is set for the twelfth, and I am to appear on that day, then I do so . . . And since all or most obligations are conditional in the sense just indicated--apply to persons in virtue of their possession of certain features, or the truth of certain descriptions of their acts and circumstances--the "options"-model appears to be applicable across-the-board.[2]

Yet Greenspan is well aware that, as we have shown, the unrestricted formula $O(p \supset q)$ will not do as a symbolization of conditional obligations. She proceeds to examine unacceptable inferences which rely on and thus appear to discredit an unrestricted form of the factual detachment rule. It might seem that if there is an obligation to see to it that q given that p, then if p is true, then one has an obligation to see to it that q. For example, if one has an obligation to show up in court if one swears to, then the fact that one swears to appears to validate the conclusion that one ought to show up in court. However, Greenspan argues that the truth of p alone is insufficient for detachment of the obligation that q. For suppose that next month I shall receive a parking ticket. Then it seems that it ought to be that if I get a ticket then I pay a fine. Presumably we would like to detach the unconditional obligation to pay the fine at least next month. But to do so now, before I receive the parking ticket, seems wrong. For my obligation now may be to avoid getting a ticket and avoid paying a fine. The truth of the antecedent, that I will in fact get a ticket, does not seem sufficient ground for detaching an unconditional obligation statement that I ought to pay a fine. Rather, Greenspan argues that such obligations do not arise until it is too late to keep their conditions from being fulfilled. Furthermore, they are no longer in force when it is too late to see to it that their objects are fulfilled. Thus when q is obligatory given p, according to Greenspan we need the "unalterability" of p, not merely the truth of p, to allow detachment of Oq. The statement that p is unalterable means, for Greenspan, unalterable for the agent at a certain time.

The proposal that conditions unalterable by an agent are required for detachment of absolute obligations motivates a non-traditional reinterpretation of the deontic operator O. Greenspan suggests that we interpret O not as indicating "what ought to be the case" but as giving directions for choice, that is, as designating what ought to be brought about. This reinterpretation of O requires the introduction of an agent relativization as well. Greenspan supposes that any particular obligation statement or argument containing several such statements will express obligations that a single given nonempty set of agents ought to bring about.[3] Thus she suggests we interpret $O(p \supset q)$ as indicating that this given nonempty set of subjects ought to bring it about that if p occurs then q does. Given this reinterpretation it is natural to defend a version of Kant's "ought implies can" principle. For what would be the point of directing an agent that he ought to bring it about that he avoid getting a ticket once he already has one? And once the condition of getting a ticket is unalterable it seems quite right to direct him to pay a fine, since avoiding getting a ticket is no longer a viable alternative. Hence Greenspan proposes the following rules as acceptable:[4]

(I) a factual detachment rule,

$$\frac{\begin{array}{l}O(p \supset q)\\ Up\end{array}}{Oq}$$

where Up asserts p's unalterable truth (unalterable by the agent) and where the statements of unalterability and obligation all hold at the same time, and

(II) a modified Kantian principle,

$$Op \supset {\sim}Up \;\&\; {\sim}U{\sim}p$$

where antecedent and consequent hold at the same time.

Greenspan asserts, then, that we may detach Oq from $O(p \supset q)$ (i) by the deontic detachment rule if Op holds at the same time that $O(p \supset q)$ holds or (ii) by the factual detachment rule (I) if Up holds at the same time that $O(p \supset q)$ holds. She concludes that $O(p \supset q)$, read "the (given) agent(s) ought to bring it about that if p then q," is applicable to conditional obligation statements as long as its application is restricted by principles (I) and (II).[5] Furthermore she argues that these rules "get us out of Chisholm's paradox"[6] and other contrary-to-duty puzzles. I shall argue that both these claims are mistaken.

According to Greenspan the time limits imposed by her "time-bound" view of obligation need not be explicit but are actually built into her principles (I) and (II).[7] If this is so, then Greenspan's description of her proposal may be inaccurate. She has not merely imposed temporal restrictions on obligation statements but has introduced a new modal operator, unalterability.

The concept of p's unalterable truth by an agent s at a time t is left undefined by Greenspan. It seems plausible to suppose that she means

(D_1) p is unalterable for s at t $=_{df}$ s does not have it in his power at t to bring about $\sim$p.

Unfortunately the concept of "having it in one's power" may be no clearer than the concept of unalterability. But if the concept of an agent having something in his power is an assertion about what is physically possible for him then we might say

(D_2) p is unalterable for s at t $=_{df}$ it is not physically possible at t for s to bring about $\sim$p.[8]

Apparently any time after p occurs we shall have not only the truth of p but the unalterable truth of p as well. But before p occurs we may have Up for s at t or $\sim$Up for s at t. If the former, then it is not physically possible at t for s to bring about $\sim p$ although it may be logically possible for $\sim p$ to occur. And in such cases the concept of unalterability and Greenspan's use of it in principle (I) appear to be most interesting.

For example, suppose that any time before Friday we may say that Smith ought to bring it about that if he accepts a job by Friday then he arrives for work on Monday. Further, suppose that during the week Smith mails a letter accepting the job. If we assume that Smith can not reach a telephone or any other means to cancel his acceptance, and cannot prevent the letter from reaching its destination, then we may say that even when no acceptance has reached Smith's future employers, still Smith's acceptance of the offer is unalterable. At such times, according to (I), we may conclude that Smith ought to bring it about that he arrives for work on Monday. Clearly (I) allows detachment of more obligations than the deontic detachment rule and fewer than the unrestricted factual detachment rule.

Principle (I) also allows the following inference. Suppose that on Monday, for example, we may say that Jones ought to bring it about that if it rains Tuesday afternoon then he closes the windows Tuesday morning. Suppose we also agree that it is unalterable for Jones on Monday that it will rain on Tuesday afternoon. Then we may conclude on Monday that Jones ought to close the windows Tuesday morning. Principle (I), with its use of the concept of unalterability, is attractive because it allows acceptable detachment of overriding obligations in cases where no detachment was possible in SDL.

The principle, however, is clearly unacceptable. According to principle (I), $O(p \supset q)$ entails Oq if p's truth is unalterable at the time the ought statements are in force. Now it seems reasonable to grant, for example, that a given agent s ought to bring it about that if he charges a $10 shirt at Macy's then he pays Macy's $10. Suppose further that s does charge a $10 shirt at Macy's, so the condition is unalterable. Then, according to (I) we can conclude that s ought to pay Macy's $10, where this is an absolute or overriding obligation. Yet this conclusion is too strong. For even if it is unalterably true that s has charged the shirt, a further condition might be sufficient to override his obligation to pay Macy's for the shirt.

Greenspan proposes an entailment substitution rule,[9]

$$\text{(III)} \quad Op \,\&\, (\sim U(p \,\&\, r) \,\&\, \sim U(\sim p \,\&\, \sim r)) \supset O(p \vee r).$$ [10]

And given this rule we shall be able to conclude that on her view if $O(p \supset q)$ holds then $O((p \,\&\, r) \supset q)$ follows even in cases where $\sim O((p \,\&\, r) \supset q)$. To see this, suppose that we grant the situation described above. S has charged a

\$10 shirt at Macy's and s ought to bring it about that if he charges the shirt then he pays Macy's \$10, which we may express by the formula $O(c \supset p)$. Suppose further that r is the state of affairs that s returns the shirt. Then as long as it is still possible for s to return or not return the shirt we may derive $O(\sim c \vee \sim r \vee p)$, or equivalently, $O(\sim(c \,\&\, r) \vee p)$ and thus $O((c \,\&\, r) \supset p)$, that s ought to bring it about that if he charges the shirt and returns it then he pays Macy's \$10. Surely this is wrong. Utilization of the concept of unalterability does not block invalid inferences from $O(p \supset q)$ to $O((p \,\&\, r) \supset q)$.

Greenspan's reinterpreted formalization $O(p \supset q)$, combined with her principles, is not applicable across the board to conditional obligation statements which are meant to assert what one really ought to do given certain states of affairs. However successful she is at avoiding the paradoxes, it is clear that her expression for conditional obligation, $O(p \supset q)$, reinterpreted as an ought-to-do, not an ought-to-be, captures some notion other than conditional obligation as explained above. It can only express a notion for which $O(p \supset q) \supset O((p \,\&\, r) \supset q)$ <u>does</u> hold, with certain restrictions.[11]

It is also clear that Greenspan has not provided an adequate way out of Chisholm's paradox. She tells us,

> No contradiction can result from a time-bound ought-to-do version of his premises, even if their conditional oughts are both detachable, as he assumes. With 'O' read as "he ought to bring it about that," and 'U' as "He cannot alter the fact that," his premises now should be understood as follows:
>
> (1) O(he goes to the assistance of his neighbors).
>
> (2) O(he goes to their assistance ⊃ he tells them he is coming).
>
> (3) O(he does not go to their assistance ⊃ he does not tell them he is coming).
>
> (4) U(he does not go to their assistance).
>
> . . . Whenever it is open to the agent to bring about the compound state of affairs, going-and-telling, that is what he ought to do, in preference to bringing about any object that includes not telling. At such times then we can detach a prescription of the consequent in (2), but not in (3). On the other hand, once his failure to go is beyond the agent's control, he ought to focus his efforts on not telling, instead. So at such times, we can detach a prescription of the consequent in (3), but not in (2). Hence (2) and (3) can never yield conflicting conclusions, but only conclusions restricted to different times.[12]

Greenspan's explanation of her solution to Chisholm's paradox may be misleading. If we let g be a statement variable for "he goes to the assistance of his neighbors" and t for "he tells them he is coming," then Greenspan's symbolization of Chisholm's sentences is:

(1a) Og

(2a) $O(g \supset t)$

(3b) $O(\sim g \supset \sim t)$

(4b) $U(\sim g)$.

Then (1a) and (2a) apparently yield Ot by the deontic detachment rule and (3b) and (4b) lead to O(∼t) by principle (I), making it difficult to see how this solves Chisholm's puzzle.

However, Greenspan's detachment rules require that we determine when the obligation and unalterability statements hold. Thus it appears that her view is something much more plausible, perhaps the following. Suppose that some particular time, say Wednesday at 3:00, is when the given agent will or will not go to assist his neighbors. Then before Wednesday at 3:00 we can say

(1a) Og
(2a) O(g ⊃ t)
(3b) O(∼g ⊃ ∼t)
(4a) ∼g.

He will not go, but it is not yet unalterable that he will not. These will not be inconsistent because (1a) and (2a) will yield Ot by the deontic detachment rule (since both hold at the same time, in this case, before Wednesday at 3:00) and yet we cannot detach O(∼t) from (3b) and (4a). However after Wednesday at 3:00, when he does not go, we shall have

(1a) Og
(2a) O(g ⊃ t)
(3b) O(∼g ⊃ ∼t)

(4b) U(~g).

In this case the four statements are inconsistent, since according to principle (II) Og ⊃ ~U(~g), and rightly so on Greenspan's view. For one can only be obligated to bring about states of affairs which are not unalterable. Presumably Greenspan would say that after Wednesday at 3:00, (1a) is false. We cannot derive Ot, therefore, but O(~t) follows from (3b) and (4b) and principle (I). Thus Ot holds only before Wednesday at 3:00 and O(~t) only after that time. On this interpretation, the conclusions are "restricted to different times" and the inconsistency pointed out by Chisholm is avoided.

As a matter of fact, however, Greenspan must maintain that on her view we can never derive O(~t).[13] For on any reasonable interpretation of Chisholm's case U(~g) holds only after he does not go. Thus O(~t), if detachable at all, is only detachable at those times after he does not go. But at such times detachment of O(~t) violates principle (II), since O(~t) ⊃ ~U(t) & ~U(~t), and at the times in question we can not have both ~U(t) and ~U(~t). After Wednesday at 3:00 it is too late to tell or not tell his neighbors he is coming. How then can Greenspan block detachment of O(~t) after Wednesday at 3:00? By pointing out that at such times (3b), O(~g ⊃ ~t), no longer holds. This is so since by (II) $O(\sim g \supset \sim t)$ implies both $\sim U(\sim t)$ and $\sim U(t)$,[14] and we have seen that after he does not go we cannot have both. Thus Greenspan must maintain that Ot may be derived before Wednesday at 3:00 but that $O(\sim t)$ may never be derived.[15] Still, Chisholm's paradox, the derivation of inconsistent obligations, is blocked.

Promising as this solution sounds, it will not do for two reasons. First, Greenspan's symbolizations do not preserve the independence of Chisholm's four sentences. Whenever it is possible for the agent not to tell his neighbors he is coming, Og will imply $O(\sim g \supset \sim t)$ by Greenspan's principle (III), and thus (1a) $\supset$ (3b). Indeed, at such times it will also be possible to derive Og $\supset$ $O(\sim g \supset t)$. Yet at no time does his obligation go to his neighbors' assistance imply a conditional obligation to tell or not to tell if he does not go. The general difficulty is that for any state of affairs r which it is still possible for the agent to bring about at the time Og holds, Og will imply $O(\sim g \supset r)$, and if the latter is to be an expression for conditional obligation, as Greenspan claims, then Og implies that the agent has a conditional obligation to anything it is still possible for him to do. Appeal to unalterability does not help Greenspan avoid this unwelcome result.

Second, Greenspan restricts her discussion to arguments in which the given obligation statements are taken to be made at the same time. This restriction prevents her from providing a consistent and insightful formulation of Chisholm's paradox. For as we have seen, her statements (1) - (4) (see quotation p. 11) cannot be held at the same time. Given principle (II) we can never assert Og and U(∿g) as true premises in the same argument, and for this reason Greenspan cannot formulate any contrary-to-duty situation consistently. This may be acceptable to Greenspan but I believe it is unhelpful. For Chisholm's point is that his four sentences may be intuitively consistent when valuated at the same time. The interesting question is what our obligations are when we have neglected our duty. Thus we want to formulate Chisholm's sentences consistently so that we may understand that an agent does have an obligation to bring about p and perhaps q, but that once he fails to do p, it is obligatory for him to see to it that ∿q.

Greenspan has claimed to present a time-bound view of obligation statements, but in addition has supplemented SDL by the introduction of a new modal operator, unalterability, and suitable principles governing detachment of obligations. Despite the plausibility of her point that detachment of obligations from avoidable conditions is acceptable, her proposal is clearly unsatisfactory. It may be, however, that another way of temporally restricting obligation statements can accomplish what we wish. Greenspan did not provide a way of designating the times at which particular obligation statements are binding. Thus our next project shall be to make a specific proposal about relativizing ought statements to times, hoping that once we can specify the time at which an obligation is in effect and the time at which it is violated, we will have a natural and non-paradoxical way of expressing obligations which arise as a result of past violations. Our aim is to determine whether temporal relativizations provide a theory which allows us to escape the contrary-to-duty paradoxes while providing a consistent symbolization of Chisholm's four sentences as independent statements and also providing an analysis of a concept of conditional obligation which will allow conditional obligations to change as conditions vary.

Let us first suppose that time can be partitioned into small units of unspecified length, which we may designate by $t_1, t_2, t_3, \ldots, t_n \ldots$ Second, let us retain Greenspan's reading of the deontic operator O as expressing what an agent ought to do. While some obligations may be in effect for a single one of our units of time, it is likely that most obligations will stand for a span of time including many of these time units. Hence it would be appropriate to designate nonempty sets of consecutive units of time, say by $T_1, T_2, T_3. . ., T_N, . . .$ during which obligations will be binding. Nevertheless, our exposition will be simplified and the points we make will be equally forceful if we restrict ourselves to expressions of obligations which hold for single units of time.

Thus we might begin by introducing the following symbolism:

$$O_{s,t_i}(p)$$

to be read "s ought (has the obligation) at t_i to bring it about that p," where s names an agent, t_i one of our time units, and p is a schematic sentence letter for a possible state of affairs.

With this symbolism we may make the following type of distinction. Suppose, for example, that Smith borrows $25 from Jones at noon on Monday. Then if p stands for paying Jones $25, and t_1 for some time on Tuesday, and s for Smith, we may say

$$O_{s,t_1}(p).$$

And if Jones, kind fellow that he is, releases Smith from his $25 debt on Wednesday, then if t_2 names some time on Thursday, we may say

$$\sim O_{s,t_2}(p).$$

This enriched symbolization is still not adequate for our purposes as it stands, however. Recall that one of our goals is to be able to distinguish conditional obligations, in particular obligations arising from neglect of our duties, from absolute obligations. Thus we shall need to temporally relativize not only obligation statements, but also our possible states of affairs, in order to be able to designate the times at which contrary-to-duty obligations arise.

For example, if Smith robs Jones, although he ought not, then Smith's contrary-to-duty obligation to make amends to Jones arises only after the robbery. Designation of a time at which Smith's contrary-to-duty obligation is binding depends upon and is relative to the time of the robbery.

Now it is important to note that conditional obligations often arise from particular actions, as in an obligation to return a book if one has borrowed it. But conditional obligations may also arise from possible states of affairs which are not actions. For example, on the condition that lightning strikes a neighbor's house and sets it on fire, one has an obligation to help the neighbor. Still, it may sometimes be difficult to designate single units of time at which possible states of affairs may be said to hold. Moreover, all our arguments will apply even if we restrict our attention to acts. So let us focus on possible states of affairs which are acts and agree that we may later augment our language to include expressions for all possible states of affairs if we wish.

We would like, then, to have a formal language within which we can symbolize sentences about acts brought about by particular agents at particular times. Suppose we let s and x range over agents named by s',s",s"',. . . Furthermore, let R and Q be predicate variables for two-place predicate letters A,B,C,. . . which we shall use to form symbolic sentences expressing relations between agents and units of time. Then we may say, for example, if s' names Smith and A_{s,t_i} expresses "s helps a neighbor at t_i" and t_1 stands for noon on Friday, that

$$A_{s',t_1}$$

will serve as a symbolization for the statement that Smith helps a neighbor at noon on Friday. In general, we shall say that any formula of the form R_{s,t_i} will be a well formed expression of our formal language.

Moreover, any formula of the form

$$O_{s,t_i}(R_{x,t_j})$$

will also be a well formed formula of our language, and such a formula will express an obligation had by an agent at a time.[16] We may read it (rather awkwardly) as asserting that s has at t_i the obligation to see to it that x bears the relation R to t_j.[17] Now we shall be able to symbolize Smith's obligation at noon on Friday to see to it that Jones helps a neighbor at 1:00 on Friday as

$$O_{s',t_1}(A_{s'',t_2})$$

where s' names Smith, s" names Jones, t_1 stands for noon Friday, t_2 for 1:00 Friday, and where A_{s,t_i} is our symbolic expression for "s helps a neighbor at t_i".

We may define the rest of our well formed formulas inductively by saying

(i) if ϕ is a wff then $\sim\phi$ is a wff,

and

(ii) if ϕ and ψ are wffs, then $\phi \mathbin{\&} \psi$, $\phi \vee \psi$, $\phi \supset \psi$, and $\phi \equiv \psi$ are wffs.

The obligation statements we are formalizing indicate what an agent ought to do, and so should yield prescriptions about what it is possible for an agent to do, in some sense of "possible." Often it may not even be physically possible for an agent to see to some obligations. For example, it may be that it is not physically possible for Smith to see to it at any time that Jones help a neighbor at that time or any other time. Jones' actions may be beyond Smith's control. So if we wish our obligation statements to yield genuine prescriptions about what an agent ought to do, it seems reasonable to begin by restricting our attention to those formulas for which (for any s,x,t_i,t_j,R)

$$(1) \quad O_{s,t_i}(R_{x,t_j}) \supset s = x.$$

Moreover, given that we can now make explicit both the time at which an ought judgment holds and the time at which the specific obligation ought to be fulfilled, it seems reasonable to specify the relationship between the two times. Although statements about what _ought to be_ may be timelessly true, in general statements about what an agent _ought to do_ will not be. It will at the very least be odd to require, for example, that at some time today Smith ought to see to it that he repays Jones yesterday. Let us formalize this intuition by adopting the minimal condition that (for any s,t_i,t_j,R)

$$(2) \quad O_{s,t_i}(R_{s,t_j}) \supset t_i \leq t_j$$

where $t_i \leq t_j$ is read "the time unit t_i occurs before or at the same time as t_j." Similarly, for statements about negative, conjunctive, and disjunctive obligations, let us require that

$$(3) \quad O_{s,t_i} \sim(R_{s,t_j}) \supset t_i \leq t_j,$$

$$(4) \quad O_{s,t_i}(R_{s,t_j} \;\&\; Q_{s,t_k}) \supset t_i \leq t_j \;\&\; t_i \leq t_k,$$

and

$$(5) \quad O_{s,t_i}(R_{s,t_j} \vee Q_{s,t_k}) \supset t_i \leq t_j \vee t_i \leq t_k.$$

Furthermore, it is reasonable to maintain that

$$(6) \quad O_{s,t_i}(R_{s,t_j}) \equiv \sim P_{s,t_i} \sim (R_{s,t_j})$$

where P_{s,t_i} is read "it is permissible for s at t_i", and where

$$(7) \quad P_{s,t_i}(R_{s,t_j}) \supset t_i \leq t_j.$$

Permissions of negations, conjunctions, and disjunctions will be formalizable analogously to (3), (4), and (5).

Let us also accept the following time relativized axioms for well formed formulas of our system, while continuing to maintain the temporal restrictions we have given:

$$(A1) \quad O_{s,t_i}(\phi) \supset \sim O_{s,t_i} \sim (\phi)$$

$$(A2) \quad O_{s,t_i}(\phi \supset \psi) \;\&\; O_{s,t_i}(\phi) \supset O_{s,t_i}(\psi)$$

and

$$(A3) \quad O_{s,t_i}(\phi \vee \sim\phi).$$

Our rule of interference, (R1), will be modus ponens.

Now, given this formal language, how are we to formalize conditional obligations? It appears that a time relativized version of $p \supset Oq$ is doomed. For we shall be able to prove $\sim R_{s,t_i} \supset (R_{s,t_i} \supset O_{s,t_j}(Q_{s,t_k}))$. And if $R_{s,t_i} \supset O_{s,t_j}(Q_{s,t_k})$ is to be a symbolization for conditional obligations, the old paradox recurs in a time relativized form. For whatever relation R that s does not bear to t_i yields a conditional obligation for s to bear <u>any</u> relation Q to <u>any</u> time t_k such that $t_j \leq t_k$. In the contrary-to-duty case, if for some s, t_m, and t_i, where $t_m \leq t_i$, $O_{s,t_m} \sim(R_{s,t_i})$ holds and yet R_{s,t_i}, then $O_{s,t_j}(Q_{s,t_k})$ follows for any s, t_j and t_k where $t_j \leq t_k$, even if $\sim O_{s,t_j}(Q_{s,t_k})$ is true.

Can a time relativized version of $O(p \supset q)$ fare any better? I think it can when we understand some natural time restrictions that may be imposed on this type of compound obligation for expressions of conditional obligation.

In the standard system of deontic logic, it was the provability of $O\sim p \supset O(p \supset q)$ which made us view $O(p \supset q)$ as inadequate as a formalization of conditional obligation (see page ix). And given the time relativized axioms (A1) -

(A3) it appears that in our time relativized deontic logic we will be able to prove

$$(8) \quad O_{s,t_i} \sim(R_{s,t_j}) \supset O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k}).$$

If so then we will indeed find $O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k})$ to be unsatisfactory as a formalization of conditional obligation. For it seems that then if s has an obligation at t_i to bear the relation $\sim R$ to t_j ($t_i \leq t_j$) then it will follow that whenever the forbidden, R_{s,t_j}, does hold, s has a conditional obligation at t_i to bear <u>any</u> relation Q to <u>any</u> time t_k later than t_i.

However, it seems to me that we can and should deny (8). We can show it will not follow when $O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k})$ symbolizes conditional obligations once we recognize and make explicit some appealing and implicit time restrictions on those obligations. For it is natural to say that if s has some absolute obligation then any conditional obligation arising from or generated by fulfillment of this obligation or its negation can only be incurred by s after this obligation or its negation is discharged. For example, I have an obligation today not to rob Smith tomorrow. Once I have robbed Smith tomorrow, then after that time I have the conditional obligation to repay him. Thus we may say that a formula of the form $O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k})$ symbolizes a sentence expressing a conditional obligation provided

$$(9) \quad O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k}) \supset t_j < t_i \leq t_k,$$

that is, only if t_i, the time of the obligation statement, is strictly after the time t_j that the condition holds. Our enriched symbolism, with this added restriction, gives us a way of making clear the time relationships implicit in conditional obligations. And every formula in our time relativized deontic logic with any implication for conditional obligations must not only follow from our axioms and principles but accord with the stipulated time restrictions as well. Surely it is now clear that (8) violates the latter. $O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k})$, as a symbolization for conditional obligation, will not follow from $O_{s,t_i}{\sim}(R_{s,t_j})$, for in the latter it must be that $t_i \leq t_j$ and in the former we must have $t_j < t_i$, which is impossible if $t_i \leq t_j$.

It appears then that the imposition of natural time restrictions has given us a formalization for conditional obligations which is immune to one of the basic difficulties that arises with an analogous, non-time-relativized and standard deontic logic. Let us now investigate whether or not $O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k})$, subject to the restrictions that $t_j < t_i \leq t_k$, is susceptible to other contrary-to-duty problems and paradoxes.

The type of contrary-to-duty puzzle discussed by Chisholm in "Contrary-to-Duty Imperatives and Deontic Logic," and Åqvist in "A Note on Commitment," has been presented in the following general way. It seems that for some possible states of affairs p and q, the following four statement forms may be consistent:

(10) It ought to be that p.

(11) Given p it ought to be that q.

(12) Given $\sim$p it ought to be that $\sim$q.

(13) $\sim$p.

Yet as we have seen, formalizations of these four English sentences in standard deontic logic lead to contradictions or entailments that seem incorrect or both. And even if we rephrase (10) - (13) in terms of what an agent ought to do we still seem to have four English sentences which may be logically consistent and mutually independent.

For example, suppose that no one should swim in the reservoir supplying town water and thus Smith ought not swim in the reservoir. But suppose Smith does swim in the reservoir and that anyone who swims in the reservoir is then required by a court order to spend a certain number of hours cleaning the grounds surrounding the reservoir. However, if Smith did not swim in the reservoir and receive this penalty he ought not to have cleaned up the grounds since it would be inappropriate to trespass on town property and others were hired to do that cleaning. Then it would be reasonable to assert each of the following four sentences:

(14) Smith ought not to swim in the reservoir.

(15) Given that Smith swims in the reservoir,

he ought to clean the surrounding grounds.

(16) Given that Smith does not swim in the reservoir, he ought not clean the surrounding grounds.

(17) Smith swims in the reservoir.

We wish to show that with our time relativized deontic logic we can symbolize these four sentences (a) as consistent statements and (b) as pairwise independent statements.

Let us suggest the following as an adequate formalization of (14) - (17), where W_{s,t_i} symbolizes "s swims in the reservoir at t_i", C_{s,t_i} symbolizes "s cleans the reservoir grounds at t_i", s' names Smith, and t_1 names the time at which Smith in fact swims in the reservoir:

(14') $(t_i)(t_j)(t_i < t_1 \;\&\; t_i \leq t_j \supset O_{s',t_i} \sim(W_{s',t_j}))$

(15') $(t_i)(t_j)(t_k)(t_j < t_i \leq t_k \supset$

$O_{s',t_i}(W_{s',t_j} \supset C_{s',t_k}))$

(16') $(t_i)(t_j)(t_k)(t_j < t_i \leq t_k \supset$

$O_{s',t_i}(\sim W_{s',t_j} \supset \sim C_{s',t_k}))$

(17') W_{s',t_1}.

It seems that there is an implicit assumption in (17) that Smith swims in the reservoir at a particular time, which we may designate by t_1, the name of one unit of time. However, although we may infer by universal instantiation from (14')

$$(18)\quad (t_i)(t_i < t_1 \supset O_{s',t_i} \sim(W_{s',t_1})),$$

that Smith ought not at any t_i earlier than t_1 swim in the reservoir at t_1, the time he does swim, the more general obligation expressed by (14') appears to be true as well and I suggest that that is what (14) asserts. In general, at <u>any</u> time before the swimming takes place we wish to assert that Smith ought not swim at any future time. The conditional obligations expressed in (15') and (16') I also take to be general or standing obligations of a sort. It is not merely true that after Smith in fact swims he must clean the grounds around the reservoir. We may also interpret (15) as claiming that at <u>any</u> time t_i, if Smith has been swimming in the reservoir at an earlier time he then ought to clean the grounds. And we may interpret (16) as indicating that at <u>any</u> time Smith ought not clean the grounds if he has not been swimming at an earlier time.

If I am correct that (14') - (17') adequately express and symbolize (14) - (17), then we do have a symbolization in which none of the four statements implies any other. For the specific occurrence of Smith's swimming at t_1 does not entail the general conditional obligation in effect any time after a swimming for Smith to clean the grounds. Thus (17') does not imply (15'). Nor will (14') imply (16') since in (14') $t_i \leq t_j$ and in (16') we require that $t_j < t_i$. We have preserved the independence of (14) - (17).

Furthermore, (14') - (17') are mutually consistent. We may, as was indicated above, infer from (14') that

$$(18) \quad (t_i)(t_i < t_1 \supset O_{s',t_i} \sim(W_{s',t_1})).$$

Similarly we may use universal instantiation to infer from (15') that

$$(19) \quad (t_i)(t_k)(t_1 < t_i \leq t_k \supset O_{s',t_i}(W_{s',t_1} \supset C_{s',t_k}))$$

and from (16') that

$$(20) \quad (t_i)(t_k)(t_1 < t_i \leq t_k \supset O_{s',t_i}(\sim W_{s',t_1} \supset \sim C_{s',t_k})).$$

But unless we wish to defend a factual detachment rule according to which $O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k})$ & R_{s,t_j} imply $O_{s,t_i}(Q_{s,t_k})$, and I think we should not, we can not infer $O_{s',t_i}(C_{s',t_k})$ from (17') and (19). And it does not appear that any time relativized version of the deontic detachment rule Op and $O(p \supset q)$ implies Oq would apply to (18) and (20) since the relationship between t_i and t_l in (18) is inconsistent with that in (20).

We have proposed a symbolization for conditional obligations which is attractive because when restricted temporally as we have suggested, it is immune to some standard paradoxes of the analogous symbolization in SDL, and looks like it will give us a way of avoiding the Chisholm-type contrary-to-duty imperative puzzle.

Unfortunately, however, there are versions of the latter puzzle which our symbolization still cannot handle. First, in proposing temporal restrictions we have also reinterpreted the deontic operator O from an ought-to-be to an ought-to-do. But this reinterpretation is so narrow that it does not allow symbolization of versions of the puzzle like the following:

(21) Jones ought not to rob Brown.

(22) Given that Jones does rob Brown, it ought to be that he be punished for robbing Brown.

(23) Given that Jones does not rob Brown, it ought to be that he not be punished

for robbing Brown.

(24) Jones robs Brown.

It may be that (22) does not express a contrary-to-duty imperative because it does not tell Jones what he ought to do once he has neglected his duty and robbed Brown. Certainly it does not tell us that Jones ought to see to it that he himself is punished for the robbery. Rather, (22) tells us what ought to be the case once Jones has robbed Brown. Thus we might not worry about this version of the puzzle nor feel the need for a solution for it because our concern is only with those obligation statements most crucial for normative ethics, that is, those which prescribe action for a given agent. Nevertheless, this version of the puzzle presents a reasonable and consistent set of deontic sentences which we cannot symbolize adequately because we have no provision for combining our time restrictions with statements about what ought to be.

Second, and perhaps more devastating to our proposal, our symbolization appears to be inadequate for Chisholm's own example and others like it. Chisholm's example, unlike the example just discussed, might be transformed into our ought-to-do language as follows:

(1) A certain man ought to see to it that he go to the assistance of his neighbors.

(2) Given that he does go he ought to see to it that he tell them he is coming.

(3) Given that he does not go he ought to see to it that he not tell them he is coming.

(4) He does not go.

Suppose we let s' name our man and let G_{s,t_i} be our symbolic expression for "s goes to the assistance of his neighbors at t_i", and let T_{s,t_i} symbolize "he tells his neighbors he is coming at t_i." Suppose also that if our man goes to assist his neighbors, there is a particular time at which he goes. Then we may let t_1 designate the time at which he would go if he were to, and may say that he does not go to help them at t_1. Then we might try to formalize (1) - (4) as

(1c) $(t_1)(t_j)(t_i < t_1 \ \& \ t_i \leq t_j \supset$

$O_{s',t_i}(G_{s',t_j}))$

(2c) $(t_i)(t_j)(t_k)(t_j < t_i \leq t_k \supset$

$O_{s',t_i}(G_{s',t_j} \supset T_{s',t_k}))$

(3c) $(t_i)(t_j)(t_k)(t_j < t_i \leq t_k \supset$

$O_{s',t_i}(\sim G_{s',t_j} \supset \sim T_{s',t_k}))$

(4c) $\sim G_{s',t_1}$.

But a closer look at (2c) and (3c) shows that the very temporal restrictions which appeared so natural and useful are devastating in this example. For normally we understand (2) to mean that if he does go he ought to tell them he is coming before he goes. Thus we require that $t_k \leq t_j$. Yet we have already restricted our time relativized version of $O(p \supset q)$ for sentences of conditional obligation so that in (2c) $t_j < t_i \leq t_k$ and hence $t_j < t_k$. Thus the restrictions we presented which were so appealing are strong enough to block some unwelcome results but too strong to allow a symbolization of all four sentences in Chisholm's very own example.

The difficulty we have met does not arise solely with Chisholm's example. For we are similarly unable to symbolize the following sentences which we might want to assert if a certain man, Smith, ought to go to New York City and can only get there by plane.

(25) Smith ought to go to New York City.

(26) Given that he goes, he ought to buy a plane ticket.

(27) Given that he does not go, he ought not buy a plane ticket.

(28) He does not go.

Normally we understand (26) to mean that Smith ought to buy the ticket before he goes. But in general our symbolic apparatus can not provide a formalization of conditional obligations such as Chisholm's (2) and (26) above, where a given condition gives rise to an earlier obligation. Perhaps we should take another look at statements like these to determine whether or not they are genuine conditional obligation statements, that is, whether or not obligations can be in effect due to future conditions, If so, then we have not solved all the puzzles of contrary-to-duty imperatives.

Finally, although a time relativized version of $O(p \supset q) \supset O((p \& r) \supset q)$ will not always hold, it still appears that as long as the relevant times are related properly we will be able to prove some invalid inferences in our time relativized deontic logic. Specifically we shall have

$$(29)\quad (t_i)(t_j)(t_k)(t_m)(t_i < t_j \leq t_k \; \& \; t_i < t_m \leq t_k \supset O_{s,t_i}(R_{s,t_j} \supset Q_{s,t_k}) \supset O_{s,t_i}((R_{s,t_j} \& S_{s,t_m}) \supset Q_{s,t_k})).$$

Neither Greenspan's use of the concept of unalterability nor our proposed time relativizations can completely block inferences of this form. Those supplements to SDL do not change most inferences allowed by use of the material conditional. It seems that we must seek a symbolization for conditional obligation statements which does not rely on the standard material implication.

NOTES TO CHAPTER I

1. Although the factual detachment rule, O(p ⊃ q) and p implies Oq is formally invalid in SDL, Greenspan is not the first to suggest that it be maintained in some form. See Wilfrid Sellars, "Reflections on Contrary-to-Duty Imperatives," Noûs 1 (December, 1967), especially pp. 305-306.

2. Greenspan, "Derived Obligation: Some Paradoxes Escaped," pp. 3-4.

3. Greenspan, "Conditional Oughts and Hypothetical Imperatives," p. 263.

4. Ibid., p. 265.

5. In response to an earlier draft of this chapter, Greenspan has replied that she did not intend to enter the debate over the proper formalization for conditional obligation sentences and in particular no longer wishes to defend her use of the material conditional connective. Unfortunately it did seem clear in both her dissertation and Journal of Philosophy paper that she was defending O(p ⊃ q) as an adequate and non-paradoxical expression when suitably restricted, and Castañeda and others have understood her to be doing so as well. (See Casteñada's "Ought, Time, and the Deontic Paradoxes," Journal of Philosophy 74 (December, 1977).) However in her most recent paper on the topic, "Practical Reasoning and Deontic Logic: Some Footnotes in Reply to Castañeda," she has dropped the material conditional connective and inserted the ambiguous English words "if...then" in its place. I commend that move and hope this dissertation contributes to the attempt to provide an adequate theory of conditional obligation which interprets the ambiguous phrase "if...then." Moreover, it may be that an acceptable theory of conditional obligation will provide further support for Greenspan's point that the truth of avoidable conditions is insufficient for detachment of unconditional obligations. If so, I would be pleased, since I am sympathetic with this view of Greenspan's in contrast to a rival view defended by Holly Goldman. (See Goldman's "Dated Rightness and Moral Imperfection," Philosopical Review 85 (1976) and Greenspan's "Oughts and Determinism: A Response to Goldman," Philosophical Review 87 (1978).) Nevertheless, I do not believe Greenspan has

provided an adequate time-bound view of oughts nor a satisfactory solution to Chisholm's paradox and the other deontic puzzles discussed in "Derived Obligation: Some Paradoxes Escaped."

6. Greenspan, "Conditional Oughts and Hypothetical Imperatives," p. 265.

7. Ibid., p. 265.

8. In correspondence Greenspan has suggested that she would prefer an analysis of unalterability which would be more acceptable to a compatibilist. For some more comments on this notion see her "Wiggins on Historical Inevitability and Incompatibilism," Philosophical Studies 29 (1976), pp. 235-247.

9. Greenspan, "Derived Obligation: Some Paradoxes Escaped," p. 53.

10. Greenspan presents the rule using the notation of her dissertation. There she says that if the time span throughout which it is possible for s to bring about $p \vee r$ is contained in the time span during which the obligation statements are in force, then $Op \supset O(p \vee r)$. More generally she says that the implication holds if it is still possible for the agent to bring about $p \vee r$ at the time the obligation statements are in force. Since she does not appeal to the rule in her more recent article she does not rewrite it in terms of U. For consistency and clarity I have attempted to do so, and I do not believe I have misrepresented her view. The point of the following example holds for either statement of the rule.

Interestingly, Greenspan does not propose a version of the more general entailment rule $(Op \,\&\, (p \supset q)) \supset Oq$.

11. This is not to say that the notion Greenspan is working with is not important or interesting. But Greenspan has indicated (in correspondence) that on her interpretation of $O(p \supset q)$, p is a sufficient condition for the conditional obligation to q. With respect to the Macy's case she says, "even given the charging of the shirt the obligation depends on some sort of admission that various 'background conditions on its force' (cf. "Conditional Oughts," p. 272) are satisfied--e.g. that s doesn't return the shirt within a certain time-span . . . p, in the Macy's case, must really stand for (or be supported by) a more complicated

condition than the one that stands out as primary, and is most convenient to focus on. . ." Given this interpretation it seems fair to characterize her notion as doubly conditional obligation. For O(p ⊃ q) indicates a conditional obligation for q given p, other things being equal, or provided there are no overriding considerations to the contrary. The difficulty is that this amounts to saying that q is conditionally obligatory given p--unless it is not. And while many general moral judgments have these ceteris paribus clauses, they do not provide us with very effective guidance. (These comments are an adaptation of those articulated by A.R. Anderson about a concept of conditional permission presented by Nicholas Rescher. (See "Reply to Mr. Rescher," Philosophical Studies 13 (1662), p. 8.)

12. Greenspan, "Conditional Oughts and Hypothetical Imperatives," pp. 265-266.

13. Greenspan does point out that there may be cases when the conditional ought in (3b) may be undetachable. (Cf. ibid., p. 266.) In what follows I show why it is always undetachable on her view.

14. This follows because O(∿g ⊃ ∿t), equivalently, O(g v ∿t), implies (a) ∿U(g v ∿t) and (b) ∿U∿(g v ∿t), that is, ∿U(∿g & t). According to Greenspan (a) yields ∿U(∿t) and (b), in combination with U(∿g) yields ∿U(t). See ibid., p. 271 and note 11, pp. 271-272.

15. Greenspan notes that an objector might find this conclusion, and her principle (I) in particular, too strong. He might wish to be able to detach O(∿t) at some time to express a contrary-to-duty obligation, perhaps even before the agent does not go so that there is still time for the agent to prevent a bad situation from becoming worse. Greenspan's response to this objector is ad hoc and unhelpful. Rather than detailing why, I shall only say that she need only have responded to her objector that directions to agents in contrary-to-duty situations may never be appropriately expressed as absolute obligations. Contrary-to-duty obligations are truly conditional, conditional upon neglect of duty.

16. It is unlikely that every two-place predicate expressing a relation between persons and times will be proper objects of obligation and I have provided no

criterion for distinguishing those that are. However, I have not attempted to define objects of obligation, but rather have proposed a symbolic language which will allow us to formalize certain English sentences expressing obligations.

17. Richard Montague, in "On the Nature of Certain Philosophical Entities," also suggests that "obligations can probably best be regarded as the same sort of thing as tasks and experiences, that is, as relations-in-intensions [predicates] between persons and moments; for instance, the obligation to give Smith a horse can be identified with the predicate expressed by 'x gives Smith a horse at t'." He goes on, however, to propose an explication somewhat different from the one presented here. He says, ". . .to say that x has the obligation R. . . would amount to the assertion that it is obligatory at t that x bear the relation-in-intension R to some moment equal or subsequent to t." p. 162.

CHAPTER II

Many of the difficulties that arise when we attempt to symbolize statements of conditional obligation by means of the formulas $O(p \supset q)$ or $p \supset Oq$ may be traced to the material conditional connective. The connective appears to be too weak for our purposes since it allows inferences we wish to block if these conditional expressions are to symbolize sentences of conditional obligation. And we have now seen that the imposition of temporal restrictions on obligation statements will not be sufficient to block those invalid inferences. There is clearly a real need for a stronger conditional expression for conditional obligation sentences.

Brian Chellas has argued in "Conditional Obligation" in Logical Theory and Semantic Analysis, edited by Stenlund, that the best approach for understanding conditional obligation divorces questions of obligation and conditionality. For then the relationship between the notions of obligation in conditional and non-conditional contexts and between the notions of conditionality in deontic and non-deontic contexts may be made explicit. Richmond Thomason argues similarly in an unpublished manuscript "Deontic Logic as Founded on Tense Logic", that a proper theory of conditional obligation will separate a theory of the conditional and a theory of obligation. For the present we shall be following their suggestion, focusing first on the notion of conditionality.

Surely entailment or strict implication, written $p \Rightarrow q$ and interpreted in the usual way, provides a stronger conditional connective than material implication. Thus we might consider utilizing it in a formalization of conditional obligation.[1] Analogs of $O(p \supset q)$ and $p \supset Oq$ provide natural candidates, and hence we shall first consider the plausibility of

(30) $p \Rightarrow Oq$

as a way of expressing conditional obligation statements.

It is probably simplest to treat the expression Oq as an "unbroken", "unanalysable" statement variable standing for English sentences of the form "It ought to be the case that q". Given this analysis, $p \Rightarrow Oq$ will be a formal expression for sentences of the form "It is necessarily the case that given p, it ought to be that q". And $p \Rightarrow Oq$ may be treated as a well formed formula of the standard alethic modal systems, where p and Oq are statement variables in those systems, since the symbol O will not function as a predicate letter or operator.

Then even in the weakest modal system T[2] we shall not be able to derive

(31) $\sim p \supset (p \Rightarrow Oq)$,

which is as we would wish. However

(32) $\Box\sim p \supset (p \Rightarrow r)$

holds for any r,[3] and so in particular

(33) $\Box\sim p \supset (p \Rightarrow Oq)$

holds for any q. On this analysis, then, an impossible condition generates a conditional obligation to anything. It is difficult to have definite intuitions about what obligations (if any) arise given impossible conditions. And so we might not be terribly troubled by (33).

But

(34) $(p \Rightarrow q) \supset (p \,\&\, r \Rightarrow q)$

is valid in T[4] and so in particular we shall have

(35) $(p \Rightarrow Oq) \supset (p \,\&\, r \Rightarrow Oq)$.

And hence if $p \Rightarrow Oq$ is to be a symbolization for conditional obligation it will fail to provide expressions for conditional obligations which change when conditions change.

Perhaps $p \Rightarrow Oq$ would be a more plausible symbolization for conditional obligation if we altered our analysis of it. Instead of treating Oq as a simple statement variable, we might take p and q to be statement variables for possible states of affairs and O to be a deontic operator read "it is obligatory that". But then $p \Rightarrow Oq$ can no longer be viewed as a well formed formula of the alethic modal logic system T. However we might augment the system T in the following way to form a formal system within which $p \Rightarrow Oq$ is a well formed expression.

We might propose that the language of our new system, call it OT, be the language of the propositional calculus, PC, plus the deontic operator O and the modal operator $\Box$. Then we might describe the well formed formulas of OT. Single statement variables will be atomic formulas of OT. And for any wffs ϕ and ψ of OT,

$$\sim\phi$$

$$O\phi$$

$$\Box\phi$$

$$\phi \;\&\; \psi$$

$$\phi \vee \psi$$

$$\phi \supset \psi$$

and $$\phi \equiv \psi$$

will be wffs of OT. This is a very rich language, for it allows mixed formulas, that is, formulas containing deontic and/or modal components as well as non-deontic, non-modal components (for example $\Box(p \;\&\; \sim Op)$) and iterated modal formulas (such as $\Box Op$). (Contrast SDL, see Appendix I.) This is just as we wish, of course, for $p \Rightarrow Oq$, equivalently $\Box(p \supset Oq)$, will certainly be a well formed formula of OT.

It seems reasonable that the axioms of OT should at the very least include all the axioms of T and SDL. Thus we might propose

(OT1) $\Box p \supset p$

(OT2) $\Box(p \supset q) \supset (\Box p \supset \Box q)$

(OT3) $Op \supset \sim O \sim p$

(OT4) $O(p \supset q) \;\&\; Op \supset Oq$.

Then in addition to the rules of PC and a rule of replacement (OT5) according to which the result of uniformly replacing any variable in any formula provable in OT by any wff of OT is itself provable in OT, we shall have

(OT6) If a formula f is provable in OT, then Of is provable in OT,

and

(OT7) If a formula ϕ is provable in OT, then $\Box\phi$ is provable in OT.

But even given this minimal description of the formal system OT, in which O functions as a deontic operator, we find that whether or not OT is objectionable on other grounds,

(36) $\Box{\sim}p \supset (p \Rightarrow Oq)$

and

(37) $(p \Rightarrow Oq) \supset (p \;\&\; r \Rightarrow Oq)$

will still be derivable in OT by applications of (OT5), (OT6), (OT7), and (OT2).[5] $p \Rightarrow Oq$ is unsuitable even when analysed as a formula of the augmented language OT.

Furthermore, regardless of the formal system within which we analyse the formula $p \Rightarrow Oq$, it appears that we shall not be able to use $p \Rightarrow Oq$ to symbolize conditional obligation statements with factual conditions. For it would seem that a factual condition cannot <u>entail</u> a normative statement. And if p expresses a non-normative state of affairs and Oq expresses an obligation that q, it is difficult to see how the former could <u>necessarily</u> imply the latter. Surely we would not wish to say, for example, that it is a necessary truth that when a promise is made then there is an obligation to keep the promise, for it is possible to make a promise that ought not to have been made, such as a promise to murder, or to make a promise and be released from that promise, and so on. In such cases the obligation to keep the promise will not follow. And thus the symbolization $p \Rightarrow Oq$, which asserts a necessary connection between a condition and the resulting obligation is too strong an expression for what seem to be paradigm cases of conditional obligation statements.

If, on the other hand, we consider

(38) $O(p \Rightarrow q)$

as our formalization for conditional obligation, where p symbolizes the condition and q symbolizes whatever is conditionally obligatory, it is difficult to understand what any conditional obligation sentence expresses. Apparently (38) says that what is obligatory is that p entail q. But suppose, for example, that I have a conditional obligation to call my sister given that I have promised to call her. Then why should it be obligatory that the relationship between my promising and my calling be a necessary one? What is obligatory is that I call my sister once I have promised to, not that there be a necessary conditional connection between my promising and doing so. It seems that (38) fails to capture the meaning of the conditional obligation sentence.

Furthermore, if we appeal again to the system OT, within which (38) is a well formed expression and O functions as a deontic operator, it is clear that

$$(39) \quad O(p \Rightarrow q) \supset O(p \;\&\; r \Rightarrow q)$$

is provable.[6] Thus (38) can only symbolize sentences for which augmented conditions do not override conditional obligations.

I have argued that neither $p \Rightarrow Oq$ nor $O(p \Rightarrow q)$ alone can be adequate as a symbolization for all the conditional obligation sentences we wish to expresss. Yet one might argue that sentences expressing conditional obligations are ambiguous and that if we used both $p \Rightarrow Oq$ and $O(p \Rightarrow q)$ we might be able to provide adequate formalizations for the sentences in question. It is unquestionable that sentences of English which we use to make statements of conditional obligation are ambiguous. For example, both the second and third sentences in Chisholm's puzzle (see Introduction, p.xii) appear to indicate conditional obligations and yet the scope of the deontic operator in each one is different. And it is an open question whether or not an adequate symbolization of them should reflect that difference or not. In other words, it is difficult to determine whether different formalizations are appropriate for the two sentences or whether the scope difference in the English sentences can be ignored and the same symbolization used for both. If different conditional obligation sentences are indeed most appropriately expressed by different formalizations, then any differences in the scope of the deontic operator in the English sentences might be a guide for determining which symbolization is suitable for a particular English sentence. But it is only a guide. There may be other features of the English sentences which would help determine the appropriate symbolization. And to single out all such features would be a nearly impossible task and beyond the scope of this dissertation. However if we admit that we may need more than one formula to express adequately the conditional obligation sentences we wish to formalize, then we must investigate whether or not $p \Rightarrow Oq$ and $O(p \Rightarrow q)$ together could be sufficient for our task.

Even without providing criteria for determining which sentences could be adequately expressed by which formulas, it seems clear that $p \Rightarrow Oq$ and $O(p \Rightarrow q)$ alone will not suffice. Neither proposal allows expression of those conditional obligations which may no longer hold given an augmented condition. Furthermore, strict implication is too strong a connective for our purposes. Any conditional obligation sentence in the form of $p \Rightarrow Oq$ indicates that there is a necessary connection between the condition and the obligation conditional upon that condition. And any sentence in the form of $O(p \Rightarrow q)$ indicates an entailment that is obligatory. And yet in the most common, paradigm, examples of conditional obligation sentences there is neither relationship. Surely there is no logical entailment between borrowing money and repaying it which makes the latter obligatory given the former. Nor is it obligatory that the relationship between borrowing and repaying be a necessary one. And thus a contingent conditional obligation to repay borrowed money, and others like it, can not be formalized by means of $p \Rightarrow Oq$ or $O(p \Rightarrow q)$. The sentences for which these formulas may be appropriate do not include the most interesting ones which may or may not be true depending on the circumstances.

Perhaps a more plausible connective to consider using in formalizing conditional obligation statements is the subjunctive conditional represented by $>$. Robert Stalnaker, in "A Theory of Conditionals" in Studies in Logical Theory, APQ Supplementary Monograph Series, pp. 98-112, first proposed and defended a theory of conditionals which provided a formal system and semantical apparatus for statements involving his conditional connective.[7] The basic conditional expression he considers, p > q, is a counterfactual statement read "if p were the case then q would be the case". Lennart Åqvist has suggested (in "A Note on Commitment", Philosophical Studies 14 (1963), p. 24) that an adequate notion of commitment may be designated by pIq, where we let "pIq denote some relation of implication that is stricter than material implication but weaker than strict implication." Stalnaker's subjunctive conditional satisfies this condition. Thus let us next consider using this connective and the semantics Stalnaker provides for it.

We might first evaluate the acceptability of

(40) p > Oq

as an expression for conditional obligation statements. If we again treat Oq as an "unbroken", "unanalysable" statement variable which serves as a formalization for sentences of the form "q is obligatory", then we might analyse (40) as a formula well formed in Stalnaker's system and read it as "if p were the case then it would be the case that q is obligatory". This proposed analysis is attractive for several reasons.

First, although $\sim p \supset (p \supset r)$ holds, $(p \supset r) \supset (p > r)$ does not in Stalnaker's system. The converse of (S9) (see note 7) is invalid. Hence we shall not be able to prove $\sim p \supset (p > r)$ and so in particular we will not have

$$(41) \quad \sim p \supset (p > Oq).$$

Thus we can avoid the paradoxical conclusion that whatever is not the case yields a conditional obligation to anything.

Second, since $p > q \supset ((p \& r) > q)$ is an invalid formula in Stalnaker's system, we will not be able to prove

$$(42) \quad (p > Oq) \supset ((p \& r) > Oq).$$

The counterfactual conditional connective apparently provides a symbolization according to which conditional obligations may change as conditions vary.

Third, although the negation of $p \supset Oq$ does not capture the negation of the conditional obligation sentence "if p then it ought to be that q", the negation of $p > Oq$ does seem to be a fair translation of the negation of "if p then it ought to be that q". For any reasonable negation of a conditional obligation sentence must assert that the obligation no longer holds under the given condition. Yet if conditional obligation statements are symbolized by means of $p \supset Oq$ then the negation of the sentence stating your conditional obligation to call your mother if she is ill is merely an assertion that your mother is ill and you are not obligated to call her. Whereas given $p > Oq$ as the symbolization for the sentence, the negation becomes the much more plausible "if your mother were ill then you would not be obligated to call her".

Fourth, if $p > Oq$, understood as a formula in Stalnaker's language, serves as our formalization for conditional obligation statements then we may have a way of formulating Chisholm's four sentences consistently and independently. If we understand both sentence (2) and sentence (3) as expressing conditional obligations, then we may write

(1a) Og

(2d) $g > Ot$

(3d) $\sim g > O(\sim t)$

(4a) $\sim g$.

As we have seen above (cf. (41)), (4a) does not imply (2d). And (1a) does not imply (3d). Neither the obligation to go nor the agent's failure to go implies that he would have any obligation whatsoever if he were to go or not to go. This formalization of the puzzle apparently yields four independent sentences.

Moreover, by (S9) (see footnote 6), (2) implies

(43) $g \supset Ot$

and (3d) implies

(44) $\sim g \supset O(\sim t)$.

And although (4a) and (44) yield $O(\sim t)$ by modus ponens, we can not derive $O(t)$ from (1a) and (43). Thus we can avoid a derivation of the inconsistent obligation statements which was possible given the formalization of Chisholm's sentences using $O(p \supset q)$ and $p \supset Oq$.

Although $p > Oq$ is thus a very plausible candidate for our project, it still appears to fall short of our expectations. First, using the subjunctive conditional connective we have given a symbolization of Chisholm's puzzle as four consistent and independent sentences. Nevertheless our symbolization is somewhat unsettling. For as was pointed out, $\sim g > O(\sim t)$ implies $\sim g \supset O(\sim t)$. Thus from (3d) and (4a) we can derive $O(\sim t)$. And this may be objectionable. For the man's obligation was to go to help his neighbors. And given fulfillment of this obligation he has an obligation to tell them he is coming.

Thus it seems odd that we can derive an absolute obligation for him not to tell them he is coming. It is especially odd if we recognize that we may derive $O(\sim t)$ only when (4a) holds. And yet once (4a) holds, it is no longer possible to discharge the obligation to tell in the sense intended in Chisholm's example. For telling or not telling presumably is to take place before going or not going. Thus given the proposed symbolization we may derive an absolute obligation not to tell when it is no longer possible to fulfill the obligation in any meaningful way.

Second, although

(45) $\sim p \supset (p > Oq)$

is not provable in Stalnaker's system, still $\Box\sim p$, or equivalently $\sim\Diamond p$, will be a sufficient condition to imply $p > Oq$.[8] That is, any impossible condition will imply a conditional obligation to anything. As mentioned above, it is difficult to have settled intuitions about whether impossible conditions can generate any obligations at all. Since one type of impossible condition, a self contradictory one, implies any consequent it may seem reasonable to allow impossible conditions to imply any conditional obligation. My own tentative intuition is that it would be preferable to have no conditional obligations follow from impossible conditions. However, we might stipulate certain changes in the formal system and semantics for the subjunctive conditional to block the conclusion that

$$(46) \quad \Box\sim p \supset (p > Oq).$$

And so I do not feel the validity of (46) in Stalnaker's system should be viewed as a reason to abandon the approach we are considering.

A third oddity which arises when using $p > Oq$ as our formalization for conditional obligation statements is that

$$(47) \quad (p \,\&\, q) \supset (p > q)$$

is provable in Stalnaker's formal system.[9] Thus in particular if condition p holds and Oq is true then $p > Oq$ follows. Any condition that is true yields a conditional obligation to any absolute obligation that holds. If Oq is true for any state of affairs q, then any fact makes it conditionally obligatory. Given that q is absolutely obligatory it may not seem objectionable to find that q is conditionally obligatory given any true condition. But it is at least odd. For then we have not captured any special relationship between the condition and the conditional obligation in such cases. And in general it does seem that there is some rather unique relationship, which we have seen can not be material or strict implication, that holds between conditions and the obligations generated by them.

A much more serious difficulty with the symbolization $p > Oq$ is evident when we examine

$$(48) \quad ((p > q) \ \& \ \sim q) \supset \sim p.$$

This formula is also derivable in Stalnaker's system. Thus we shall be able to prove

$$(49) \quad ((p > Oq) \ \& \ \sim Oq) \supset \sim p.$$

And if $p > Oq$ is to express conditional obligation statements then (49) tells us for example that if an agent has a conditional obligation to repent if he sins and has no absolute obligation to repent then it follows that he does not sin. This is clearly untenable. Surely the antecedent may be true and the consequent false. We must be able to express all of the following in some consistent manner: that a state of affairs q is not absolutely obligatory, but that q is conditionally obligatory given condition p and that p is true. However no such situation can be expressed consistently if $p > Oq$ is our formalization for conditional obligation within Stalnaker's system.

We might believe, however, that our attempt to analyse p > Oq as a formula of Stalnaker's system was misguided. Perhaps that analysis obscured important features of obligation and we would be more successful if we considered p > Oq as a formalization within which O functions as a deontic operator.

Unfortunately, however, although we have an axiomatic system governing > and an axiomatic system for O when applied to material conditional statements, no formal system has been developed which combines the connective > and the deontic operator O. There are thus no straightforward syntactic methods available for determining whether certain formulas involving both > and O, and analogous to the paradoxical formulas of SDL extended to include mixed formulas, are true or false.

However, in order to evaluate the adequacy of p > Oq, we might attempt to develop an extension of SDL which would provide the syntactic apparatus we are lacking. To begin, we might supplement the basis logic, BL, of SDL by incorporating the connective > in the language of BL, appropriate rules for forming well formed formulas, including mixed formulas, involving >, and Stalnaker's axioms governing >. Given the latter, certain other symbols will have to be incorporated in the language.[10] Thus we might simply take as our augmented basis logic Stalnaker's logic for the connective >. In Stalnaker's system validity can be defined in the usual way and he has shown his system to be complete in the sense that every valid formula is a theorem. Hence it seems satisfactory as an augmented basis logic. (Compare Hansson's description of BL, see Appendix I.)

We might form an augmented deontic logic, call it OS, from S analogously to the formation of SDL from BL, by adding the deontic operator O to the language of S, and adding wffs of the form Of where f is a formula of S. The axioms, in addition to those of S, will be the axioms of SDL. We shall have a rule of replacement allowing uniform substitution of wffs of OS for variables in formulas provable in OS to yield formulas provable in OS, and a rule according to which if a formula f is provable in S, then Of is provable in OS. Clearly $p > Oq$ will be a well formed formula of OS.

Unfortunately, even given our new formal system OS, in which O functions on its own as a deontic operator, $p > Oq$ can not be adequate for expressing all sentences of conditional obligation. For as long as we maintain Stalnaker's axioms for the subjunctive conditional, now applicable to formulas including the deontic operator O by the replacement rule, we must still accept

(49) $((p > Oq) \& \sim Oq) \supset \sim p.$[11]

Indeed, as long as we maintain that a version of (S9) holds, namely that

$$(50)\quad (p > Oq) \supset (p \supset Oq),$$

and it seems we must, it does not appear that we can avoid the provability of (49). Hence we are again left with the unwelcome result that if $p > Oq$ is to serve as a symbolization for conditional obligation, then we cannot formalize a case in which q is not absolutely obligatory, but is conditionally obligatory given condition p and that p is true. And surely such cases are common in normative ethics and are the basis for a major portion of the conditional obligation sentences we wish to formalize. I may now have no absolute obligation to take a book to the library, but it may be that if I do borrow a book from the library then I am conditionally obligated to take back the book. It is consistent with these that I do borrow a book from the library. Yet using $p > Oq$ we can not symbolize these sentences.[12]

We have noted that the subjunctive conditional is midway in strength between strict implication and material implication, that is,

$$(51)\quad (p \Rightarrow q) \supset (p > q) \supset (p \supset q)$$

holds in Stalnaker's system (and thus also in our augmented system OS.)[13] And it is the only conditional connective we have considered thus far which blocks certain inferences we wish to avoid. Given the unacceptability of $p > Oq$, then, it is natural to consider

(52 $O(p > q)$

as a formal expression for conditional obligation.

In order to analyse $O(p > q)$ we need to view it as a well formed formula of a formal language allowing expressions involving both the subjunctive conditional $>$ and the deontic operator O. It is natural to refer again to the system OS described above. We might also want to choose additional plausible looking axioms combining $>$ and O. For example, we might want to add as an axiom

(53) $O(p > q) \supset (Op > Oq)$.[14]

If $O(p > q)$ serves as our expression for conditional obligation, then according to (53), a conditional obligation for q given p implies that if p were absolutely obligatory then q would be absolutely obligatory. Thus, for example, a conditional obligation to keep a promise given that it is made implies that if it were obligatory to make the promise then it would be obligatory to keep it.

This in turn implies that if it is obligatory to make the promise then it is obligatory to keep it. If it is not obligatory to make the promise then this entire material implication is true, but that does not seem objectionable.

In a contrary-to-duty case, (53) says that a conditional obligation to repent given a sin, which we would represent by $O(s > r)$, implies $Os > Or$, that is, if it were obligatory to sin then it would be obligatory to repent. And given $Os > Or$, it will follow that $Os \supset Or$, an obligation to sin materially implies an obligation to repent. The falsity of the antecedent forces the truth of $Os \supset Or$, but since $Os \supset Or$ has no implication in terms of conditional obligation, its vacuous truth when it has a false antecedent seems unproblematic.

It is thus at least plausible to add (53) to OS. Furthermore, its addition is advantageous if $O(p > q)$ is to be our symbolization for conditional obligation since it provides a straightforward rule for detachment of unconditional obligations from conditional ones. For whenever q is conditionally obligatory given p and p itself is obligatory then we may detach an absolute obligation for q. That is,

$$(54) \quad \frac{\begin{array}{l} O(p > q) \\ Op \end{array}}{Oq}$$

will be a valid inference since $O(p > q)$ will imply $Op > Oq$ and this in turn yields $Op \supset Oq$. An application of modus ponens then justifies the conclusion. We may want to detach Oq even more often than (54) allows, but (54) provides at least a minimal detachment rule which could be supplemented.

Given the formal system OS it is appropriate to consider certain formulas analogous to those which caused difficulty for the symbolization $O(p \supset q)$. In particular, if $O(p > q)$ is our symbolization for conditional obligation we hope

(55) $O(\sim p) \supset O(p > q)$

is not derivable, since we do not want the forbidden to conditionally obligate one to anything. Given the axiomatic system as presented above we find that $O(\sim p)$ implies nothing whatsoever about the truth of $O(p > q)$. Although $O(\sim p)$ does imply $O(p \supset q)$, as we have seen, there is no implication from $O(p \supset q)$ to $O(p > q)$. An assumption of the negation of (55) does not lead to an inconsistency.

When we consider

(56) $O(p > q) \supset O((p \& r) > q)$

we find that the obligatoriness of $p > q$ implies nothing whatsoever about the obligatoriness of $(p \& r) > q$.

Furthermore, when we use $O(p > q)$ to symbolize Chisholm's four sentences, we can produce in many ways the most satisfactory formalization proposed thus far for that puzzle. Let us again suppose that both sentences (2) and (3) express conditional obligations. Then we may have

(1a) Og

(2e) $O(g > t)$

(3e) $O(\sim g > \sim t)$

(4a) $\sim g$.

As pointed out above, (1a) does not imply (3e). And it seems clear that (4a) does not imply (2e). $\sim g$ is not a sufficient condition to yield the modalized counterfactual conditional in (2e). This is as we wish since the fact that he does not go does not imply a conditional obligation to tell or not tell given that he does go. Thus we apparently have a symbolization of four independent sentences. Also (2e) implies $Og > Ot$ by (53) and this implies $Og \supset Ot$. Similarly, (3e) implies $O(\sim g) > O(\sim t)$ and this yields $O(\sim g) \supset O(\sim t)$. Thus from (1a) and (2e) we may derive $O(t)$ but we can not derive $O(\sim t)$ from (3e) and (4a), and hence we do not have the inconsistency pointed out by Chisholm. Moreover, that we can derive O(t) but not O(∿t) (as opposed to being able to derive O(∿t) but not O(t) from the symbolization on page 49) seems quite correct. For it is right to say that his obligation is to go to his neighbors' assistance, and to tell them he is coming. He has no absolute obligation not to tell them he is coming, however, although he ought not to tell given that he does not go.

It appears, then, that the formula O(p > q) provides a way of expressing sentences of conditional obligation which is immune to the difficulties which arose for any of the other symbolizations we have considered. Encouraged as we are by these findings, it is difficult to be completely satisified at this point with O(p > q). We may wonder what English sentence provides an acceptable translation of expressions in the form O(p > q). Are we to read O(p > q) as asserting that a certain counterfactual conditional is obligatory? And if so, what does that mean? Moreover, how are we to determine the truth value of substitution instances of O(p > q)? The syntactic apparatus provided by OS may show that certain inferences will not follow from the formula O(p > q). But it does not help us determine whether or not those English sentences which we believe express true conditional obligations will be true when symbolized in the form O(p > q).

In order to confirm our conclusions about $p \Rightarrow Oq$, $O(p \Rightarrow q)$, and $p > Oq$ and provide a more complete evaluation of $O(p > q)$ by discovering answers to these questions, we must seek semantic analyses for these formal expressions.

NOTES TO CHAPTER II

1. Notably, Alan Ross Anderson has proposed an analysis of "commitment" in terms of strict implication in "On the Logic of 'Commitment'", Philosophical Studies 10 (1959), pp. 23-27. He suggests that we define 'pCq' which is read "p commits us to q" or "q is obligatory given p" as $p \rightarrow Oq$ where the arrow indicates strict implication. However he adds a note indicating his debt to Kripke for pointing out that if the underlying alethic modal system is strong enough (S5) then his proposal is inadequate. (Op. cit. p. 26.) The proposal is criticized carefully by Lennart Åqvist (in "A Note on Commitment", Philosophical Studies 14 (1963), pp. 22-25) and his analogous proposal for conditional permission is criticized in Nicholas Rescher's "Conditional Permission in Deontic Logic", Philosophical Studies 13 (1962), pp. 1-6. Anderson abandoned his original proposal.

More recently G. H. von Wright proposed an analysis for commitment or conditional obligation in "Deontic Logic and the Theory of Conditions", (in Deontic Logic: Introductory and Systematic Readings, Hilpinnen, ed.). He defines a new deontic operator Q as follows:

$$Q(p/q) =_{df} Sc(p, Oq) \mathbin{\&} \sim\Box(p \supset q),$$

where Q(p/q) is read "it ought to be the case that q, given that p," (op. cit. p. 169) and where

$$Sc(p, Oq) =_{df} \Box(p \supset Oq) \mathbin{\&} \Diamond p \mathbin{\&} \Diamond\sim p.$$

Hence his proposal can be expressed as

$$Q(p/q) =_{df} \Box(p \supset Oq) \mathbin{\&} \Diamond p \mathbin{\&} \Diamond\sim p \mathbin{\&} \sim\Box(p \supset q).$$

Thus von Wright also proposes an analysis of conditional obligation in terms of strict implication.

In light of Anderson's and von Wright's work it is important to investigate the plausibility of providing an analysis of conditional obligation in terms of alethic modal logic and deontic logic. In what follows I explain the futility of such approaches for our concept. The criticisms I develop below show that each of Anderson's and von Wright's proposals are ultimately unsatisfactory for our purposes.

After abandoning his original proposal Anderson suggests that $p \rightarrow Oq$ may be adequate as a symbolization for commitment if the arrow is interpreted as entailment in the sense of his system E. ("Reply to Mr. Rescher", _op. cit._, p. 6.) In that system we can avoid the problem posed by vacuous truth of the conditional with a false antecedent. However we are still able to prove

$$(p \rightarrow Oq) \supset (p \,\&\, r \rightarrow Oq).$$

For this reason and in light of Åqvist's criticisms (_op. cit._), even this proposal will be inadequate for our concept of conditional obligation.

2. The alethic modal system T (sometimes called M), first developed by C. I. Lewis, is built from the propositional calculus (PC). The language of T is the language of PC together with the symbol $\Box$, read "it is necessarily the case that". The axioms of T are the axioms of PC as well as the following two modal axioms which can be stated in terms of statement variables p and q,

(T1) $\Box p \supset p$

and (T2) $\Box(p \supset q) \supset (\Box p \supset \Box q)$,

a replacement rule, (T3), and a rule of necessitation according to which, for any formula ϕ of T,

(T4) if $\vdash_T \phi$ then $\vdash_T \Box\phi$.

The stronger modal systems B, S4, and S5 are formed by adding additional axioms to T.

3. $\sim p \supset (p \supset r)$ is a theorem of the propositional calculus and so of the modal system T as well. Hence $\Box(\sim p \supset (p \supset r))$ is a theorem of T by (T4). Then $\Box\sim p \supset \Box(p \supset r)$ follows by (T2) and (T3).

4. By applications of (T3), (T4), and (T2).

5. With respect to (36), since $\sim p \supset (p \supset Oq)$ is derivable from the axioms of PC, (OT6), and the rule of replacement (OT5), then $\Box(\sim p \supset (p \supset Oq))$ follows by (OT7) and so $\Box(\sim p) \supset \Box(p \supset Oq)$ by (OT2) and (OT5).

For (37), $(p \supset Oq) \supset ((p \,\&\, r) \supset Oq)$ follows from the axioms of PC, (OT6), and (OT5), and so $\Box((p \supset Oq) \supset ((p \,\&\, r) \supset Oq))$ by (OT7). Hence $\Box(p \supset Oq) \supset \Box((p \,\&\, r) \supset Oq)$ by (OT2) and (OT5).

6. $\Box(p \supset q) \supset \Box((p \& r) \supset q)$ is provable in OT since it follows from the rules of PC, (OT7), (OT2) and (OT5). Thus by (OT6) and (OT5), $O(\Box(p \supset q) \supset \Box((p \& r) \supset q))$ is provable in OT, and so by (OT4) and (OT5) we have $O\,\Box(p \supset q) \supset O\,\Box((p \& r) \supset q)$, or equivalently, $O(p \Rightarrow q) \supset O((p \& r) \Rightarrow q)$.

7. In Stalnaker's formal system for his conditional connective, if p, q and r are statement variables,

(S1) $\Box p =_{df} \sim p > p$

(S2) $\Diamond p =_{df} \sim(p > \sim p)$

and (S3) $p \gtrless q =_{df} (p > q) \& (q > p)$.

The axioms of his system are:

(S4) Any tautologous well formed formula of the propositional calculus

(S5) $\Box(p \supset q) \supset (\Box p \supset \Box q)$

(S6) $\Box(p \supset q) \supset (p > q)$

(S7) $\Diamond p \supset (p > q) \supset \sim(p > \sim q)$

(S8) $p > (q \vee r) \supset (p > q) \vee (p > r)$

(S9) $(p > q) \supset (p \supset q)$

(S10) $p \gtrless q \supset (p > r) \supset (q > r)$,

where $\Box$ and $\Diamond$ are the necessity and possibility operators. The inference rules are

(S11) modus ponens

and (S12) the rule of necessitation. (See (T4), note 1.)

David Lewis later provided a similar analysis of subjunctive conditionals in Counterfactuals, Harvard University Press, 1973.

8. For $\sim p \supset (p \supset r)$ is an axiom for any r by (S4). Thus we have $\Box(\sim p \supset (p \supset r))$ by the rule of necessitation and so $\Box(\sim p) \supset \Box(p \supset r)$ by (S5). Since $\Box(p \supset r) \supset (p > r)$ by (S6) we have $\Box(\sim p) \supset (p > r)$ for any r. With Oq substituted for r, $\Box(\sim p) \supset (p > Oq)$.

9. For suppose we had $p \& q$ and $\sim(p > q)$. Then we would have $p \& q$ and $\sim(p \supset q)$ by (S9) and hence $(p \& q)$ and $(p \& \sim q)$, which is impossible.

10. E.g., $\Box$, $\Diamond$, cf. (S5), (S6), in note 6, for example.

11. (49) follows from (S4), (S9) and the replacement rule.

12. It is now clear that we meet the same difficulty with $p \Rightarrow Oq$. Both $p > Oq$ and $p \Rightarrow Oq$ imply $p \supset Oq$. Thus if either serves as our symbolization for conditional

obligation then we must accept a general factual detachment rule, since truth of p implies truth of Oq. And as Greenspan has persuasively argued, (see Chapter I, pp. 9-10) such a rule allows detachment of too many unconditional obligations. David Lewis makes the same point, with very little explanation, on page 102 of _Counterfactuals_.

13. See p. 44, note 6, (S6) and (S9).

14. This has been suggested by Hans Lenk in "Varieties of Commitment", _Theory and Decision_ 9 (1978), pp. 17-37.

CHAPTER III

The standard deontic logic discussed thus far, with or without temporal restrictions, and the systems within which we analysed alternative conditional obligation formulas using strict implication or a subjunctive conditional, have all been presented as axiomatic systems. It is noteworthy that for several years deontic logic was viewed purely syntactically. The adequacy of an axiomatic system was tested by deriving theorems from a set of axioms, translating those theorems into sentences of English and judging the intuitive plausibility of those sentences. The difficulties of this approach are now well known. Choice of a set of axioms is often achieved merely by trial and error. Moreover, translations of formulas into English often involve ambiguous expressions such as "implies", "requires", and "commits", which make it difficult to understand the English sentences and thus to determine how intuitively acceptable their formal counterparts are as principles of deontic logic.

For these reasons the adaptation of the semantics for alethic modal logic provided by Kripke to deontic logic has been a welcome advance. Model theory provides a systematic way of interpreting deontic formulas and judging their acceptability based on their assigned meanings.

In Chapter II we found that $p \Rightarrow Oq$, $O(p \Rightarrow q)$, and $p > Oq$ were unsatisfactory formalizations for conditional obligation and that $O(p > q)$ appeared to be more acceptable when analysed in the syntactic system OS. However it is difficult and perhaps impossible to evaluate the acceptability of the system OS within which we can formalize $O(p > q)$, without any semantic apparatus. For many complex formulas of OS, such as substitution instances of the proposed axioms, it is difficult to find a meaningful and understandable translation into English with which to evaluate their acceptability as axioms in an augmented deontic logic. And without any semantic apparatus we have no systematic way of comparing alternative axiomatic bases. Furthermore, to determine the truth value of instances of $O(p > q)$ we need a semantic interpretation of that formula. Thus it is essential that we provide semantic analyses for the formalizations we are considering to confirm the unacceptability of $p \Rightarrow Oq$, $O(p \Rightarrow q)$, and $p > Oq$, and to complete our evaluation of $O(p > q)$.

Jaakko Hintikka was one of the first of several modal logicians to present semantical theories for deontic logic. Moreover, in "Some Main Problems of Deontic Logic", he addresses in particular what he calls the problem posed by the notion of commitment. That is, he asks what can be meant by saying that a certain fact or act p commits one to acting in a certain way, q.[1]

The phrase "conditional obligation" has often been used to refer to any obligation arising from any state of affairs, whether that state of affairs is an act or not. And the term "commitment" usually refers in particular to those conditional obligations which an agent incurs based on his or her own actions. Clearly, Hintikka does not make use of this conventional distinction. Indeed, his description of what he calls the problem of the notion of commitment indicates that his focus is what we would call the problem of conditional obligation. Hintikka suggests that his model theoretical interpretation of deontic logic shows that the traditional formalizations in deontic logic can be used to express at least some commitments or conditional obligations without being paradoxical. I shall argue to the contrary that $O(p \supset q)$ and $p \supset Oq$ are inappropriate for expressing conditional obligation or commitment sentences even when interpreted with Hintikka's semantics. However I go on to appeal to his semantic interpretation to complete the evaluation of $p \Rightarrow Oq$, $O(p \Rightarrow q)$, $p > Oq$, and $O(p > q)$ as symbolizations for conditional obligation sentences.

For Hintikka, statements about what is obligatory or permitted are taken to be statements not about what people actually do or fail to do, but about what they do and should do in a possible world where all obligations are fulfilled. These expressions are thus counterfactual and contain an implicit reference to alternative worlds. The basic notion in Hintikka's semantics for deontic logic, then, is the notion of a deontic alternative to a given world, a kind of "deontically perfect world". A deontically perfect alternative to a world, the actual world, for example, is defined by Hintikka as a world in which all obligations obtaining in the actual world, as well as any further obligations obtaining in the deontic alternatives, are fulfilled. Moreover, according to Hintikka, whatever is permissible in the real world obtains in a deontic alternative. Since not all permissions in the real world can be made use of in the same world, (as when p and not-p are both permissible in the actual world), we must consider more than one deontic alternative to a given world. In particular, if w_2 is a deontic alternative to w_1 Hintikka proposes the following conditions:

(i) If Op is true at w_1 then p is true at w_2,

(ii) If Op is true at w_2 then p is true at w_2,

and (iii) If Op is true at w_1 then Op is true at w_2.

We may focus on (ii) and (iii) since taken together they imply (i). Furthermore, according to Hintikka,

(iv) (Pp is true at $w_1 \supset (\exists w_i)$ (w_i is a deontic alternative to w_1 and p is true at w_i),

and (v) Op is true at $w_1 \supset (\exists w_i)$ (w_i is a deontic alternative to w_1 and p is true at w_i).

Hintikka's description of deontic alternatives as deontically perfect worlds is a problematic one and we shall return to discuss it in more detail. However if we suppose for the present that we can understand and use the notion of a deontically perfect alternative to a world, we may examine Hintikka's claim that a world theoretic interpretation provides a firm foundation for traditional symbolizations for conditional obligation.

Hintikka makes a distinction between logical consequence and deontic consequence. q is a logical consequence of p just in case $p \supset q$ holds, that is, just in case p and $\sim q$ are not both true, whereas q is a deontic consequence of p just in case $O(p \supset q)$ holds, that is, just in case p and $\sim q$ are not both true in any deontically perfect alternative. Given this distinction, it is clear that the two candidates in standard deontic logic (extended to include mixed formulas) for formulating commitments or contional obligations,

(57) $O(p \supset q)$

and (58) $p \supset Oq$

are not equivalent.

Rather than focusing on one of these, Hintikka believes it is an inescapable conclusion that "our commonplace notion of commitment is intrinsically ambiguous between the two renderings"[2] and possibly still others. Thus he maintains that some sentences expressing commitments and conditional obligations may be formalized using $O(p \supset q)$ while others are best symbolized by $p \supset Oq$. He believes that his semantic interpretation shows the paradoxes of derived obligation to be "but particular cases of the paradoxes of implication, and hence devoid of special interest for a student of deontic logic."[3]

According to Hintikka,

(59) $O(\sim p) \supset O(p \supset q)$

and (60) $Op \supset O(q \supset p)$

are indeed valid, however,

> if the validity of [59] and [60] is looked upon from the point of view of our 'deontically perfect worlds', the appearance of a paradox is considerably dimished. In [59], it is true to say that p cannot be realized in a deontically perfect world without realizing q because p cannot be so realized simpliciter. In [60], q cannot be realized in a deontically perfect world without realizing p, for p has to be realized in any such perfect world in the first place. Thus the 'paradoxes' lose their sting against our interpretation [$O(p \supset q)$], provided we realize what precisely it contains. At worst we have a residual feeling of awkwardness which can be traced to the same sources as the usual 'paradoxes' of entailment (implication). 4

Similarly,

(61) $\sim p \supset (p \supset Oq)$

is valid in Hintikka's system. Yet he claims that "what creates the appearance of paradox here is not so much the idea on which [61] is based as rather the desire to have some stronger implicational tie between p and Oq than a material implication..."[5] And whatever the motivation for desiring a stronger connective, pragmatic or otherwise, Hintikka nevertheless claims that there are many deontic notions for which p ⊃ Oq, with material implication, serves well.

The world theory interpretations Hintikka gives for (57) and (58) do provide one way of understanding those formulas, as well as (59), (60) and (61), to be non-paradoxical. However the interpretation does not tie those formulas to any concept of conditional obligation or commitment. And if (57) and (58) are to serve as symbolizations for commitment, then it seems quite clear that Hintikka can not deny that by (59) a forbidden act commits one to everything, by (60) everything commits one to an obligatory act, and by (61) the realization of whatever is not in fact realized commits one to everything. And these are the paradoxes Hintikka claims his interpretation avoids. Hintikka's interpretation does show that on that interpretation the relevant formulas are unproblematic. But his interpretation alone does not justify using those formulas for symbolizations of conditional obligation or commitment. In fact, given his interpretation, it is difficult to understand why (57) and (58) are even likely candidates for symbolizations of commitment.

Moreover, Hintikka's semantic analysis does not block the derivation of the inferences we have discussed which must be invalid if (57) and (58) are to serve as expressions of commitment. For if p logically implies Oq then p & r logically implies Oq. If p and ∼Oq can not both be true, then p & r and ∼Oq can not both be true. And if q is a deontic consequence of p, then q is a deontic consequence of p & r. That is, if p and ∼q are not true in any deontically perfect world then p & r and ∼q can not be true in any of those worlds. Thus on Hintikka's view if either (57) or (58) is a symbolization for commitment, then if p commits one to q, then p & r commits one to q for <u>any</u> q. On his view changing conditions do not give rise to changing commitments.

It is clear that Hintikka's interpretation, though it may be helpful in assessing the acceptability of deontic statements, does not alone transform (57) and (58) into satisfactory symbolizations for commitment. Hintikka himself notes in particular that if p is forbidden then in such "unusual circumstances" (57) will not be useful for the purpose at hand. For if p is forbidden we can not have any way of determining what p implies in a deontically perfect world; since no transgressions occur in deontically perfect worlds, p can not occur in such a world at all. But surely it is not at all uncommon for agents to do the forbidden in the actual world. And we have stressed that an adequate theory of normative ethics will direct agents to act in certain ways given their transgressions. These directions, or contrary-to-duty imperatives, are what we should like our deontic logic to be able to express. Whatever other deontic notions (57) and (58) may express adequately, they will not do for expressions of commitment or conditional obligation, particularly contrary-to-duty obligations, even when given Hintikka's model theoretic interpretation.

Although Hintikka's semantic analysis fails to transform $O(p \supset q)$ and $p \supset Oq$ into suitable formalizations for the sentences we wish to symbolize, nevertheless it may help us evaluate the other expressions we have considered. I think it can be used to confirm the conclusions of Chapter II that $p \Rightarrow Oq$, $O(p \Rightarrow q)$, and $p > Oq$ can not serve as the symbolizations we desire.

First, we noted that any conditional obligation sentence in the form $p \Rightarrow Oq$ would indicate a necessary connection between the condition and the conditional obligation. And since it seems clear that factual conditions can not entail normative statements, $p \Rightarrow Oq$ is a poor symbolization for conditional obligation sentences. If we apply the Kripke semantics for alethic modal logic and Hintikka's semantics for deontic formulas according to which

(62) Op is true at w_i ≡ p is true at every deontic alternative (deontically perfect world with respect) to w_i,

in the most natural way, we can confirm these results.

For consider

(63) Oq ⊃ □Oq,

a formula in our system OT. And suppose we are evaluating this formula with respect to w_r, the real world. Then we might have Oq true at w_r, and thus q true at all the deontic alternatives to the real world, although □Oq is false at w_r. For an evaluation of □Oq presumably requires determining the truth value of Oq at <u>every</u> world which is an alternative to w_r. And even though q is true at all deontic alternatives to w_r, there is no assurance that q will be true at all deontic alternatives to <u>every</u> alternative of the real world. Thus if $w_1 \neq w_r$ is an alternative to w_r, then q may be true at all the deontic alternatives of w_r but false at some deontically perfect world with respect to w_1.

Since (63) will not hold, then even if Oq is true, we will not always be able to affirm

(30) p ⇒ Oq.

For (30) can be interpreted as saying that in <u>all</u> worlds w_i, if p is true at w_i then Oq is true at w_i, that is, q is true at all deontically perfect worlds with respect to w_i. But as we have seen, even if Oq is true at the deontic alternatives to a given world where p holds, it does not follow that q holds at all deontic alternatives of <u>every</u> alternative where p holds. Thus if p is true at w_1 and $w_2 \neq w_1$ is an alternative of w_1 where p is true, then q may be true at all the deontic alternatives of w_1 but false at some deontically perfect world with respect to w_2.

For example, it may be that in w_1 Smith has received a manuscript from Jones and writes Jones promising to return it. Then presumably in w_1 Smith ought to return it, and so in deontically perfect worlds with respect to w_1, Smith sends the manuscript back to Jones. It also may be that in w_2, Smith, believing that he has been sent the manuscript, writes to Jones promising to return it. However suppose that in w_2 Smith never did have the manuscript; Jones still has the paper. Then there is no reason to believe that Smith keeps his promise in the deontic alternatives to w_2, and thus ought to return the manuscript in w_2, even though he has made the promise in w_2. Presumably it is physically impossible for him to return the manuscript in the deontic alternatives of w_2. And although Hintikka does not specify all the relevant details about deontically perfect alternatives, it is plausible to assume that in such worlds agents do not perform acts they are physically unable to perform. But more generally, in this case failure to keep the promise does not appear to be at all incompatible with deontic perfection. Hence we can not conclude that Smith keeps the promise in all the deontic alternatives of w_2.[6]

With respect to $O(p \Rightarrow q)$, a semantic analysis using the Kripke semantics for strict implication and Hintikka's deontic semantics in the most obvious way shows that $O(p \Rightarrow q)$ is true at the real world just in case $p \Rightarrow q$ is true at all the deontically perfect worlds with respect to w_r. But $p \Rightarrow q$ is true in some deontic alternative to w_r, w_1 say, just in case $p \supset q$ is true in all the alternatives to w_1. If we assume we are evaluating the formula in S_5 so that every world is an alternative to every other world, and agree that deontic alternatives are possible worlds, then $O(p \Rightarrow q)$ is true just in case $p \supset q$ is true at every world, that is, just in case $p \Rightarrow q$ holds. And surely any symbolization for <u>any</u> type of obligation sentence, conditional or otherwise, which implies $Op \supset p$ is unacceptable.

And finally, when we consider $p > Oq$ as our formalization we again confirm our conclusion that it is unacceptable. For when we interpret $p > Oq$ using the basic Hintikka-type semantics for the deontic operator O and Stalnaker's semantics for the subjunctive conditional, we find that there are many conditional obligation sentences we believe to be true that we can not express formally as true sentences. For applying Stalnaker's semantics first, to evaluate the truth value of p > Oq (assuming p is possible), we must consider the world closest or most similar to the actual world where p holds, w_1 say, and determine whether or not Oq holds there. If so, p > Oq is true, and if not, p > Oq is false. In order to evaluate the truth of Oq at w_1 we must again appeal to the definition assigning truth values to deontic formulas according to which Op is true at a world just in case p is true at every deontic alternative to that world. Then Oq will be true at w_1, the world most similar to the real world where p is true, just in case q is true at all the deontic alternatives of w_1.

But we can see that, for example, we can not express a true conditional obligation for an agent to repent if he sins. For if we let s > Or express Smith's conditional obligation to repent if he sins, then this expression can never be true. We must first consider the possible world most similar to the actual world in which Smith sins, and then determine whether or not Smith repents in all the deontic alternatives to that closest world in which he sins.

When we do this we inescapably face some of the difficulties of the standard Hintikka-type deontic semantics. We might wonder, for example, whether Smith will even exist in the deontic alternatives to the real world. Apparently on Hintikka's analysis if Smith has any obligations at all then he will exist in deontic alternatives to see to their fulfillment. That is, if Op is true then p is fulfilled in all the deontic alternatives. Hence if fulfillment of p entails Smith's existence then Smith will exist in all the deontic alternatives. Nevertheless, we might ask whether Smith exists in those alternatives as the same type of person as he is in the actual world. If, in the real world, he is the type of person who never makes reparation for past injustices, will this be compatible with deontic ideality? Suppose we imagine that the deontic semantics are restricted (although presently they are not) in such a way that Smith exists in deontic alternatives as much like himself as is consistent with the deontic ideality of those worlds.

Then, given that on the deontic semantic analysis we are utilizing the deontic alternatives are "deontically perfect worlds" with respect to the world where Smith sins, it is clear that it is not true at all those worlds that Smith repents. Indeed, Smith does not sin or repent at any of those deontically perfect worlds. Thus s > Or is false in this example.[7]

And p > Oq will be false given Stalnaker's and Hintikka's semantic interpretations whenever the conditional obligation is incompatible with deontic ideality. Thus in Chisholm's puzzle we may express an agent's conditional obligation to tell his neighbors he is coming given that he goes to their assistance, but we can not express his conditional obligation not to tell if he does not go as a true conditional obligation sentence using the symbolization $\sim g > O(\sim t)$. Given the semantical interpretations for O and > that we have been considering, that formula will always be false. The agent does not fail to tell in the relevant deontic alternatives. And hence the symbolization we proposed on page 49 presented Chisholm's four sentences as formally consistent and independent, but not all true when analyzed using these semantic theories.

It has often been pointed out that a major difficulty with Stalnaker's semantics for the counterfactual conditional is the problem of determining a similarity ordering of possible worlds in order to establish the closest world in which the condition holds. However the problem we have just discussed does not turn on the ambiguity of the similarity relation. Rather, the basic deontic world theory analysis, with its appeal to deontically perfect alternatives, is at fault. We must conclude that $p > Oq$ is unsatisfactory as an expression for conditional obligation statements if it is interpreted semantically as above. And we have seen that given Stalnaker's axiomatic system for the subjunctive conditional it is formally inadequate as well.

Given Hintikka's deontic interpretation we have been able to reaffirm our view that $p \Rightarrow Oq$, $O(p \Rightarrow q)$, and $p > Oq$ can not serve as acceptable formalizations for conditional obligations. However, O(p > q) is not susceptible to similar formal difficulties. Nevertheless, semantic analysis indicates that there are objections to using O(p > q) as our formal expression as well.

First, if we again apply the semantic analysis provided by Stalnaker for > and Hintikka's semantics for O, if p is true in some world and q is also true in that world, then the subjunctive conditional statement, p > q, is true in that world. Thus if p and q are both true in the deontic alternatives with respect to some given world, then O(p > q) is true in that world. It appears, then, that on this analysis too much will be conditionally obligatory.

It might be replied that if p and q are both true in the deontic alternatives to some world, then given the nature of these alternatives, p and q are both consistent with deontic perfection and hence the truth of O(p > q) for such states of affairs is unobjectionable. Nevertheless, it must be admitted that O(p > q) will then express true conditional obligation statements between unobjectionable but also sometimes unrelated states of affairs.

A second and much more serious difficulty is that we can find examples which show that given these semantic analyses for > and O, conditional obligation sentences symbolized by O(p > q) may turn out to be false when we believe them to be true. For example, suppose s stands for

"Smith sins" and r for "Smith repents". On the proposed analysis we would symbolize Smith's conditional obligation to repent if he sins by $O(s > r)$. To evaluate the truth value of $O(s > r)$ we must determine whether or not $s > r$, the counterfactual conditional within the scope of the deontic operator, holds at all the deontic alternatives to the actual world. At this point the semantic analysis gets more complicated. For according to an extended version of Stalnaker's semantics for the subjunctive conditional, $s > r$ is true at some deontic alternative to the real world, w_1 say, just in case the consequent r is true at the world most similar to w_1 where the antecedent s holds. Thus we must determine the truth of r at worlds most similar to deontic alternatives of the real world.

Let us again assume that Smith exists in the deontic alternatives to the actual world as much like himself as is consistent with the deontic ideality of those worlds. Then clearly in the deontic alternatives to the actual world he neither sins nor repents. Neither is compatible with deontic perfection. But then what reason do we have or could we have for believing that in those worlds most similar to the deontic alternatives (of the real world) in which he does sin that he repents? We might be tempted to conclude that the worlds most similar to deontically ideal alternatives in which he does sin are deontically perfect except for his sinning (and any other changes required by his sinning). We might thus believe them to be the best worlds in which he sins, and conclude that in those worlds Smith does repent for his sins. However that conclusion is clearly unjustified, for it is tantamount to an assumption that similarity is measured with respect to goodness, and of course such an assumption can not be made.

It seems just as reasonable to conclude that Smith does not repent in the worlds most similar to the deontically ideal alternatives where he sins. For suppose w_1 is a deontic alternative to the real world. Then Smith neither sins nor repents in w_1. Hence a world w_2 in which Smith sins and repents may be less similar to w_1 than a world w_3 which is just like w_2 except that Smith sins and does not repent there. In any case we have no clear way of determining which of w_2 and w_3 is more similar to w_1. We can not merely count the number of properties that each of w_2 and w_3 has in common with w_1, since in both cases the number will be infinite. And any intuitive argument that one of w_2 or w_3 is more similar to w_1 than the other will apparently appeal to features about Smith as he exists in deontic alternatives that are left undetermined by Hintikka's deontic semantics. Hence we are not justified in concluding that the semantics guarantees that Smith repents in all worlds most similar to deontic alternatives where he sins. And without this guarantee we can not assert the truth of O(s > r) when it is interpreted using the above semantics for the obligation operator and the subjunctive conditional.

Clearly, an application of these semantic analyses to evaluate instances of formula O(p > q) requires two world relativizations and is extremely cumbersome. It is also, because of the ambiguity of the similarity relation required for the interpretation of the counterfactual conditional, difficult to determine the truth value of some instances of O(p > q). But it seems clear that in at least some cases, a conditional obligation sentence will be true although its formal counterpart as an instance of O(p > q) will be false.

Our discussions in this chapter have forced us to face some of the difficulties of Hintikka's deontically perfect world semantics. What is deontic perfection? Can we make sense of the notion? Is it even a consistent concept as used in Hintikka's theory? Apparently not. Hintikka tells us that all obligations of the real world will be fulfilled in its deontically perfect alternatives. But then how can "good samaritan" obligations to help criminals, for example, be fulfilled in deontic alternatives if people always behave as they ought in those worlds? Furthermore, we may ask who is in these worlds? Will each deontic alternative of a given world have the same inhabitants?

Will inhabitants of deontically perfect alternatives be the same type of people as they are in the given world? Hintikka gives us little information for answering these questions.

Moreover, as we have clearly emphasized, an analysis of deontic alternatives as deontically perfect is blatantly inappropriate for an analysis of those conditional obligations which are contrary-to-duty. For example, Smith ought not to harm Jones, but if he does, he has an obligation to repair the harm done. Yet no account at all can be given of obligations to make reparation using Hintikka's deontically perfect world semantics; acts of reparation are not fulfilled in deontically perfect alternatives and so on his analysis are not obligatory. Since a major aim of our project is to provide a symbolization and analysis which can account for contrary-to-duty obligations, we must revise Hintikka's deontic semantics.

One natural revision, which has been suggested by Føllesdal and Hilpinnen[8] and by Bengt Hansson,[9] is (roughly) to maintain the same interpretation as Hintikka's for formulas of the form Op, but to introduce a new dyadic obligation operator, O(p/r), for conditional obligation sentences. According to Føllesdal and Hilpinnen, O(p/r) "means that p is true in all (possible) worlds in which r is true, and which resemble deontically perfect worlds as much as possible."[10]

First, all of the difficulties which arise for Hintikka's theory of unconditional obligation will remain problems on this interpretation. For example there will be no consistent way of maintaining that there can be "good samaritan" obligations to those who have committed morally forbidden acts.

Second, if the notion of a deontically perfect world is problematic and vague, then the notion of an "almost deontically perfect" world is even less helpful. An extension of Hintikka's semantics in this way is not as easily managed as Føllesdal, Hilpinnen, and Hansson would have us believe. How are we to understand the description that these deontic alternatives "resemble deontically perfect worlds as much as possible" except for the truth of the condition? How extensively will these worlds differ from the given world and from deontically perfect worlds? What will these worlds be like? To emphasize the difficulties of this proposed extension of Hintikka's semantics, let us consider the following example.

Suppose Smith has a conditional obligation to buy his wife some medicine since he has lost her full medicine bottle. Suppose also that he has promised to buy a friend a book and so has a conditional obligation to buy the book. Let us suppose further that he does not have enough money to buy both the medicine and the book, although he can buy one or the other. Since he cannot fulfill both conditional obligations we may wish to deny that he has both conditional obligations. But presumably we would want to say he does have at least one of the conditional obligations, perhaps the obligation to replace the lost medicine, since it would provide the most utility, or be the most stringent prima facie duty, or whatever (depending on the moral theory we appeal to). Yet if we extend Hintikka's theory to apply to conditional obligations as proposed above, then Smith has this conditional obligation just in case he buys the medicine in all those deontic alternatives which are deontically perfect except for his loss of the medicine and any other changes required by that loss.

Is fulfillment of that conditional obligation even compatible with the degree of deontic perfection in those worlds? If we answer no, on the grounds that Smith must then break his promise and not buy the book, then we appear to have a counterexample against the proposal. Thus advocates of the proposal will have to argue that buying the medicine is compatible with the deontic perfection in these alternatives. But how can a proponent of this view support such an argument?

He might first say that Smith is richer in the relevant alternatives and so can buy both the book and the medicine. However it seems implausible that anything in the vague description of these deontic alternatives as deontically perfect except for the loss of the medicine assures us that this is so. Even if we adopted the suggestion made earlier that Hintikka's semantics be supplemented with the assumption that the agent exists in the deontic alternatives as much like himself as is consistent with the requisite deontic ideality, it seems implausible that we could be assured that Smith would be richer in the deontic alternatives merely because he has also promised the friend to buy him a book.

Second, a proponent might argue that Smith never makes the promise to buy the book in the deontic alternatives and thus it is compatible in the almost-deontically perfect worlds for him to buy the medicine. But on what grounds can he claim Smith never makes the promise in the deontic alternatives? _Making_ the promise is not incompatible with deontic perfection. It is _fulfillment_ of the promise that is incompatible with _fulfillment_ of the conditional obligation to buy the medicine. And if the advocate gives _this_ as a reason to claim that Smith never makes the promise in the relevant alternatives he is appealing to a characterization of the deontic alternatives not only modified to accomodate the truth of the condition, but also tailored in terms of what insures that fulfillment of the conditional obligation is compatible with the near-deontic perfection of the alternatives. However since the semantic analysis directs us to determine the truth of a conditional obligation based on whether or not it is fulfilled in all the deontic alternatives then to be non-circular we need a clear characterization of those alternatives independently of whatever is compatible with fulfillment of the obligation.

Our example shows that the modification in Hintikka's semantics appealing to "nearly deontically perfect worlds" presented above is much more problematic than it appears to be. I see no way of determining whether or not in the case just described Smith buys the medicine in all those worlds in which he loses the medicine and which are otherwise as deontically perfect as possible. The concept of an almost-perfect deontic alternative is too vague to provide an adequate foundation for an analysis of conditional obligation.

David Lewis has proposed a deontic semantics which differs from Hintikka's more radically than the proposal we have just discussed and thus it may be more plausible.[11] He abandons the notion of deontically perfect worlds altogether. Lewis' semantic insight is to evaluate a formula Op at a world by reference to deontic alternatives, but to characterize deontic alternatives differently. On his view we imagine that the deontic alternatives of a world can be ranked according to some unspecified moral principle or set of moral principles. Then a formula Op will be true just in case p is true at the "best" deontic alternatives. Lewis also proposes a dyadic obligation primitive for conditional obligation, accompanied by a semantic analysis in terms of "best" worlds. This type of suggestion will be teleological in that obligatoriness and rightness will be defined in terms of goodness, although the notion of goodness, which must be defined independently of rightness, may be left undetermined in the formal analysis. This approach, with an appeal to "best" rather than "deontically perfect" worlds may provide a more plausible framework within which to analyze conditional obligations in general and especially contrary-to-duty obligations.[12]

Let us pause at this point to clarify the notion of conditional obligation we are attempting to analyze. Our work thus far has shown that we are focusing on a concept for which (i) there is a non-trivial connection between conditional obligations and their conditions; on our view it will not follow that everything or nothing is conditionally obligatory, (ii) conditional obligations may no longer hold if overriding conditions hold, (iii) further conditions may reinstate (and override again, etc.) the original conditional obligation, (iv) truth of the condition alone does not justify detachment of an unconditional obligation from a conditional one, (v) contrary-to-duty obligations are a type of conditional obligation, and (vi) sentences in the form of Chisholm's four sentences, where the second and third are both taken to be conditional obligation sentences, may be truly, consistently and independently expressed at the same time.

We must conclude from our discussions in this chapter and the last, that a change in the material conditional connective alone can not transform either $O(p \supset q)$ or $p \supset Oq$ into an expression for a notion of conditional obligation which satisfies the above conditions. It appears, then, that we must develop a new obligation operator for our concept of conditional obligation.

Interestingly, David Lewis' proposal satisfies many of the above conditions. Nevertheless, satisfaction of the six conditions above does not narrow down a notion of conditional obligation sufficiently. For Lewis explicitly presents his analysis of conditional obligation as an analysis of a conditional ought-to-be. Yet even if it ought to be that there is no illness in the world, it may be that no individual ought to do anything in particular to see to it that there is no illness in the world.

Our hope is to develop a symbolization and analysis for a concept which satisfies the six conditions above (and perhaps others) and which is, moreover, a conditional ought-to-do, as opposed to a conditional ought-to-be.[13] There may be other concepts of conditional obligation satisfying different conditions from those specified above. However I believe that we have intuitions about a concept which does satisfy all those conditions and that it is <u>this</u> concept which is most important for normative ethics.

In Chapter IV we shall carefully explain and evaluate Lewis' theory, which satisfies so many of the conditions we have focused upon, although it provides an analysis of a concept of conditional ought-to-be.

NOTES TO CHAPTER III

1. Hintikka, "Some Main Problems of Deontic Logic", in Hilpinnen, ed., Deontic Logic: Introductory and Systematic Readings, Dordrecht, 1971, p. 87.

2. Ibid., p. 87.

3. Ibid., p. 88.

4. Ibid., p. 88.

5. Ibid., pp. 89-90.

6. In this example we have had to assume we have some plausible, consistent, and intuitive idea of a deontically perfect world. Yet this may not be the case. (See below, pp. 79-80,85-6.) Furthermore, the foregoing arguments rely rather heavily on Hintikka's deontic semantics, in which there is no guarantee that deontic alternatives of one world are the same as the deontic alternatives of every world. It may be that for some other deontic semantics, suitably restricted, (63) and (30) will be true. If so, then given that type of interpretation a factual condition apparently will entail Oq for some q. We shall discuss this further in Chapter IV.

7. As I point out below this analysis relies heavily on Hintikka's interpretation of a deontic alternative as a deontically perfect world. A different deontic interpretation might not provide such a telling argument against the symbolization p > Oq. Nevertheless, in view of our other criticisms of p > Oq (see pp. 50-56), I believe it is fair to judge it as unsuitable for the notion we wish to capture.

8. "Deontic Logic: An Introduction" in Deontic Logic: Introductory and Systematic Readings, Hilpinnen, ed., Dordrecht, 1971, p. 26, pp. 29ff.

9. "An Analysis of Some Deontic Logics", Hilpinnen, ed., op. cit., pp. 121-147.

10. Hilpinnen, ed., op. cit., p. 30.

11. David Lewis, Counterfactuals, Harvard University Press, 1973, chapter 5.

12. We must develop the theory more fully but the intuitive idea is that in the above example even if Smith loses the medicine and makes the promise in some deontic alternatives, he fulfills the obligation arising from the former in all the "best" alternatives in which he loses the medicine.

13. In this respect, then, we are pursuing the notion important to Greenspan. Thus we might more properly follow her by referring to "conditional oughts" rather than "conditional obligations". Nevertheless, we shall continue to use the phrase "conditional obligation" for our notion of a conditional ought-to-do. This use of "obligation" instead of "ought" for grammatical ease is not uncommon and is not intended to be misleading. Cf. Holly Goldman, "Dated Rightness and Moral Imperfection", Philosophical Review 85 (October, 1976), pp. 478-479.

CHAPTER IV

David Lewis proposes a semantic analysis for deontic logic in Chapter 5 of his recent book Counterfactuals. His analysis is based on a semantic insight which is similar to Hintikka's, namely that formulas of the form Op are true at a world just in case p is true at all the deontic alternatives of that world. However Lewis proposes a novel interpretation of deontic alternatives. He assumes that deontic alternatives of a given world may be ranked based on their comparative goodness.[1] He says,

> Suppose we have a preference ordering of the worlds, perhaps different from the standpoints of different worlds. As is the custom in deontic logic, I shall say nothing definite about the source and significance of this ordering. Perhaps the worlds are ordered according to their total net content of pleasure, measured by some hedonic calculus; or their content of beauty, truth, and love; or their content of some simple, non-natural quality. Perhaps they are ordered according to the extent that their inhabitants obey the law of God, of Nature, or of man. Perhaps according to how well they measure up to some sort of standards of objective morality, if such there be; perhaps according to the tastes we would have if we attained a superhuman capacity for calm, sympathetic, impartial contemplation of alternative possibilities. It does not matter. We can build in the same way on any of these foundations, or on others.[2]

Intuitively, we may imagine the deontic alternatives of a given world w_1 as arranged in a system of spheres of accessibility around w_1. The accessibility relation is that of "being evaluable from." Thus the deontic alternatives may be viewed as part of a system of spheres of evaluability; each world or deontic alternative in the system of spheres may be evaluated or ranked by the unspecified standards of evaluation that give rise to the ordering from the standpoint of w_1.

If the preference ordering from the standpoint of w_1 is a weak ordering, that is, is reflexive, transitive, and connected, then a world w_2 is better than a world w_3 with respect to w_1 just in case some sphere contains w_2 but not w_3. The idea is that spheres closer to the center of the system contain worlds that are better, or are ranked higher, with respect to the preference ordering. Moreover, worlds which differ only in respects which are wholly irrelevant to their comparative goodness, or for which relevant differences are balanced out, will be tied in the preference ordering, and thus will occupy comparable positions in the system of spheres.

Given this description, we might be tempted to view the system of spheres as centered or weakly centered.

> Centering means that for each world i, {i} is a sphere around i; and that would mean that each world is, from its own standpoint, the best of all possible worlds. Weak centering means that each world i belongs to the innermost nonempty sphere around i; and that would mean now

> that each world is, from its own standpoint, at least one of the best worlds.[3]

But, Lewis correctly notes,

> It is quite clear, no matter what (within reason) is the source of our preference ordering, that ours is nowhere near being one of the best possible worlds![4]

Thus the system of spheres based on comparative goodness will not be centered or weakly centered.

Lewis suggests that it is natural to impose a normality condition on the system of spheres. That is, we may wish to grant that some world is evaluable from every world, or, more clearly, that every world is such that some world is evaluable from it. This condition is reasonable since it turns out that in worlds where this condition does not hold nothing is obligatory and everything is permitted.

Lewis leaves open the question of whether or not to add a condition of universality which specifies that every world is evaluable from every other world. If we impose universality then we abandon evaluability restrictions.

Similarly, Lewis leaves open the question of whether or not to accept absoluteness. If absoluteness does hold, then whatever relation there is between two worlds with respect to one world, holds between them with respect to any other world. That is, the relative goodness of deontic alternatives will not vary from world to world but will be absolute; deontic alternatives will have the same preference ordering from the standpoint of every world. As Lewis points out, an ordering of worlds according to the amount of utility in the world will presumably be the same from the standpoint of any world. However an ordering based, for example, on the extent the inhabitants of a world obey the laws of the ruling god in the given world will differ from the standpoint of different worlds with different gods.

It is an asset of Lewis' semantic analysis for obligation statements that the criterion used for ranking the deontic alternatives is left unspecified. However if we assume that whatever criterion is used will be an objective moral standard of some kind, then it will be reasonable to accept both universality and absoluteness. Nevertheless, Lewis' formal theory does not commit us to doing so.

Lewis explicitly states that he is proposing an analysis of obligation as what ought-to-be the case.

> 'Obligation' is here used in a special, impersonal sense. What is obligatory (conditionally or unconditionally) is what ought to be the case, whether or not anyone in particular is obligated to see to it. Personal obligations may or may not follow from these impersonal obligations.[5]

One formulation of Lewis' theory about what ought to be the case is:

(64) Op (read "it is obligatory that p") is true at $w_i \equiv$
i) $(\exists w_j)$ (w_j is evaluable from the standpoint of w_i)
& ii) p holds at all worlds that are best from the standpoint of w_i.

This interpretation, however, can only be applied if Lewis' Limit Assumption holds, that is, if there is an innermost nonempty sphere relative to each world containing all the worlds which are best from the standpoint of that world. And Lewis is careful to explain that we can not assume that the Limit Assumption does hold. "We might have an infinite ascent to better and better worlds, and no innermost sphere containing best worlds of all. For every world there would be a sphere small enough to exclude it, so the intersection of all nonempty spheres would be empty."[6]

Thus, if there are no best worlds with respect to w_i, but rather an infinite ascent to better and better worlds, then Lewis proposes

(65) Op is true at $w_i \equiv$
i) $(\exists w_j)$(w_j is evaluable from

the standpoint of w_i)
& ii) p holds throughout all worlds which are sufficiently good from the standpoint of w_i.

We might rephrase the second clause more perspicuously and have

(66) Op is true at $w_i \equiv$
$(\exists w_j)\{(w_j$ is evaluable from the standpoint of w_i) &
$(w_k)(w_k$ is at least as good as w_j from the standpoint of $w_i \supset$ p is true at $w_k)\}$.

For conditional obligation sentences (also about what ought to be the case) of the form "Given that p, it ought to be the case that q," (which I shall symbolize p ⥽ q), Lewis provides an analysis of a deontic primitive analogous to (64) which depends on the Limit Assumption:

(67) p ⥽ q is true at $w_i \equiv$
i) $\sim(\exists w_j)(w_j$ is evaluable from w_i & p is true at w_j)
or ii) $(\exists w_j)(w_j$ is evaluable from w_i & p is true at w_j) & q holds at all the worlds where p holds which are best from the standpoint of w_i.

And we might reformulate (67) in a manner similar to our reformulation of (64). Thus

(68) $p \mathrel{\circ\!\!\rightarrow} q$ is true at $w_i \equiv$
i) $\sim(\exists w_j)(w_j$ is evaluable from w_i & p is true at $w_j)$
or ii) $(\exists w_j) \{(w_j$ is evaluable from w_i & p is true at $w_j)$ & $(w_k)(w_k$ is at least as good as w_j from the standpoint of w_i & p is true at $w_k \supset q$ is true at $w_k)\}$.

If we assume that since we are assessing moral obligation, universality and absoluteness do hold and hence every world is evaluable from every other world, and the relative ranking of worlds will be the same from the standpoint of every world, then we may restate (66) and (68) as

(69) Op is true $\equiv (\exists w_i) \{(p$ is true at $w_i)$ & $(w_j)(w_j$ is at least as good as $w_i \supset p$ is true at $w_j)\}$

and

(70) $p \mathrel{\circ\!\!\rightarrow} q$ is true $\equiv$
i) $\sim(\exists w_i)(p$ is true at $w_i)$
or ii) $(\exists w_i) \{(p$ is true at $w_i)$ & $(w_j)(w_j$ is at least as good as w_i & p is true at $w_j \supset q$ is true at $w_j)\}$.

It is readily apparent that if (70) is our analysis of conditional obligation then any impossible condition yields a conditional obligation to _anything_. We may recall that the same result followed for (30), $p \Rightarrow Oq$ (cf. p. 39), and for (40), $p > Oq$ (pp. 51-52). Lewis is well aware of this result, and is not sure whether an impossible condition generates all or no conditional obligations. Thus he presents two conditional obligation analyses from which we may choose. His alternative primitive for conditional obligation sentences has an analysis exactly like (70) except that the conditional obligation sentence is false rather than true whenever p is impossible no matter what q is. I find the latter preferable (see p. 51) and so shall focus on Lewis' alternative primitive, namely,

(71) $p \mathbin{\circ\!\!\Rightarrow} q$ is true $\equiv$
$(\exists w_i)\{(p$ is true at w_i & q is true at $w_i)$ & $(w_j)(w_j$ is at least as good as w_i & p is true at $w_j \supset q$ is true at $w_j)\}$.

This proposal of Lewis' for conditional obligation sentences is a particularly plausible one for several reasons. First, even if $p \mathbin{\circ\!\!\Rightarrow} q$, it will not generally follow that $p \& r \mathbin{\circ\!\!\Rightarrow} q$. That is, even if there is a world where p and q hold such that all equally or higher ranked worlds are such that both p and q hold in them, there is no guarantee that there will be a world where p & r and q hold and for which p & r and q all hold in all equally good or better worlds. It may be best for q to hold when p does, but not when both p & r do.

Second, $p \Rightarrow q$ is a primitive and is not defined in terms of the unconditional obligation operator. Thus we have no rules for factual detachment of unconditional obligations from conditional ones. And so happily we need not accept a principle justifying factual detachment of an unconditional obligation based on the truth of the condition alone. It would be appealing if we could propose some rule for relating conditional and unconditional obligations, but we may propose one; no such principle is imposed by the semantic analysis.

For this reason a symbolization of Chisholm's sentences (interpreted as sentences about what ought to be the case) as

(1a) Og

(2f) $g \Rightarrow t$

(3f) $\sim g \Rightarrow t$

and (4a) $\sim g$,

where the second and third formalizations are interpreted according to (71), is appealing. We might think that it is better for the agent to go and tell than merely to go to aid his neighbors, and hence believe that if there is a world where he goes such that he goes in all better worlds, then surely there is a world where he goes and tells such that he goes and tells in all worlds at least as good. Thus we might believe that (1a) implies (2f). However nothing in the formal semantics assures us of this. The implication depends on the particular criterion of goodness used for ranking worlds, and this criterion is clearly left unspecified.

Thus, even if there is a world where the man goes to help his neighbors which is such that he also goes in all worlds as central or more central in the system of spheres, there is no assurance that there will be a world where he goes and tells them he is coming such that he goes and tells in all worlds within the sphere containing that world. Nor will it follow that there is a world where he neither goes nor tells such that he neither goes nor tells in all equally or higher ranked worlds. And the fact that he goes to help them in the real world implies nothing whatsoever about what happens in better worlds. The four formalizations are truly independent on this formal analysis.

But is the symbolization proposed above consistent? We might be tempted to conclude that from a conditional obligation for q given p that we can detach an obligation for q. For we might think that if there is a world where both p and q are true such that in all equally good or better worlds p and q are true, then there is surely a world in which q holds such that q is true in all equally good or better worlds. But that reasoning would be mistaken. For even if given a certain world, q is true in all the better worlds where p is also true, there might be even better worlds where p is not true and neither is q. Thus $g \Rightarrow t$ does not yield Ot nor does $\sim g \Rightarrow \sim t$ yield $O(\sim t)$.

We have already noted that truth of the condition is insufficient on Lewis' view for detachment of an unconditional obligation. Thus (4a) and (3f) do not yield $O(\sim t)$. That he fails to go in the real world and that worlds in which he fails to go and fails to tell appear in the ranking in a certain order, indicate nothing whatsoever about how worlds in which he fails to tell (whether or not he goes) appear in the ordering.

Thus even if (1a) and (2f) together yield Ot we will not have the inconsistency noted by Chisholm. Nevertheless it is interesting to ask whether or not (1a) and (2f) <u>do</u>, on this view, lead to Ot. Suppose there is a world where the man goes to help his neighbors such that he goes in all worlds at least as good. And suppose further that there is a world in which he goes <u>and</u> tells such that he tells in all equally good or better worlds in which he goes. Then it follows that there is a point at which equally or higher ranked worlds are all such that he tells in them. Thus we apparently can affirm a deontic detachment principle

$$(72) \quad \frac{p \Rightarrow q \qquad Op}{Oq}$$

using Lewis' analysis. And even with this detachment principle we have found that Chisholm's sentences are not inconsistent given the formalizations above.

Another nice feature of Lewis' proposal is that we can compare deontic alternatives. Clearly proposals which classified all deontic alternatives as deontically perfect, or which distinguished perfect and near-perfect alternatives, were found to be inadequate. As Lewis says,

> . . . a mere division of worlds into the ideal and less-than-ideal will not meet our needs. We must use more complicated value structures that somehow bear information about comparisons or gradations of value.[7]

In our example from pages 87-90 we could not determine whether or not Smith bought a new bottle of medicine for his wife in all deontic alternatives which were deontically ideal except for his loss of the medicine (and other changes required by the loss). However, on Lewis' view, we need not determine whether or not certain states of affairs are compatible with some vague amount of deontic perfection. Rather, we can consider *all possible* alternatives in which he loses the medicine, including those in which he promises to buy his friend the book and has not the money to buy medicine and the book. Then he has the conditional obligation to buy the medicine if he buys it in those alternatives in which he loses it and which are sufficiently good in the *comparative* ranking of deontic alternatives. And this ranking is determined only when we *independently* superimpose a definite principle or set of principles for ordering the deontic alternatives.

It seems clear, then, that Lewis' proposal is superior to others we have examined. But it is appropriate to make some evaluative comments and also to analyze two recent criticisms of Lewis' theory of conditional obligation.

In Chapter III we noted that given Hintikka's interpretation of deontic operators

$$(63)\quad Oq \supset \Box Oq$$

does not hold. Interestingly, given Lewis' semantic analysis for obligation the truth of (63) turns on his absoluteness and universality conditions. If absoluteness does not hold, then the ordering of deontic alternatives may differ from world to world and so q might be true at all the best worlds with respect to w_r but false at some best world with respect to $w_1 \neq w_r$, an alternative of w_r. Hence, (63) will not be valid.

However, if we adopt absoluteness and universality, and maintain that the ordering of deontic alternatives is the same from the standpoint of every world, and that every world is evaluable from the standpoint of every other world, then (63) is valid and so we can see that this version of Lewis' semantic analysis is not equivalent to Hintikka's. For Oq is true at w_r just in case there is a world, w_1 say, evaluable from w_r, and such that q is true at all worlds at least as good as w_1 from the standpoint of w_r. Thus if absoluteness and universality hold, then for every world w_i it is true that there is a world evaluable from w_i, namely w_1, such that q is true at all worlds at least as good as it is.

What are the implications of the truth of (63)? It appears that

$$(29) \quad p \Rightarrow Oq$$

(alternatively, $\Box(p \supset Oq)$), will be true for *any* p if Oq is true. Thus it seems that a factual condition *can*, on Lewis' view (assuming absoluteness and universality) *entail* an obligation statement; yet we have argued that such an entailment is clearly implausible.

Although (29) is true according to Lewis' semantics for alethic and deontic logic when we assume absoluteness and universality <u>whenever Oq is true</u>, (29) is not true for any p and any q. Thus the provability of (29) in the special case when Oq is true, though perhaps odd, may be a harmless consequence of the theory.

To see this, notice first that (29) implies $p \supset Oq$. Thus if (29) is true and p is also true we may derive Oq. But since (29) is true <u>if</u> Oq is true then derivation of Oq in this instance is unsurprising.

Second, since the entailment in (29) will hold only if q is obligatory <u>per se</u>, truth of (29) in such a case merely indicates that if q is obligatory then any condition entails that it is obligatory. And although I find this result odd, it seems no more peculiar than the analogous result in alethic modal logic, that necessary truths are entailed by any factual or contingent statements, as is clear from

$$(73) \quad \Box q \supset \Box(p \supset \Box q).$$

As long as (29) has no implication in terms of conditional obligation, as we have argued it can not, then its truth when Oq is true may be harmless. And if we find the result worrisome we might block it by restricting Lewis' system (although he does not) so that iterated modalities are not allowed. This type of restriction is imposed by Bengt Hansson on his deontic systems.[8]

A more serious peculiarity of Lewis' view turns on his analysis in terms of best worlds. We have pointed out that even if p ⇏ q, it will not generally follow that p & r ⇏ q. However there may still be cases in which p ⇏ q implies p & r ⇏ q on Lewis' analysis even though we believe the English sentence corresponding to the former to be true and that corresponding to the latter to be false. For example, it seems reasonable to assert that

(74) Given that Smith promises to mail Jones an application form, then it ought to be that he mail a form to Jones.

and to deny that

(75) Given that Smith promises to mail Jones an application form and that he hand carries one to Smith, then it ought to be that he mail a form to Jones.

In other words it is reasonable to claim that the second condition, bringing a form in person, overrides the conditional obligation to mail one.

Yet it seems to me that on Lewis' view, if (74) is true then (75) must be true as well. To see this, suppose we agree that given Smith's promise it is better for him to bring Jones an application in person than to mail one, (whether or not he does mail one). The time it saves Jones outweighs any inconvenience to Smith. Then we may say that worlds in which Smith promises to mail a form, mails one, and also hand carries one to Jones, are better than worlds in which he merely promises to mail a form and does so. Thus if it is true that there is a world in which Smith promises to mail Jones an application such that all worlds at least as good in which he promises are such that he does mail a form to Jones, then it will follow that there is a world in which Smith promises to mail a form and also carries one to Jones such that in all equally good or better worlds in which both conditions hold he mails a form to Jones.

The point of the example can be generalized. If, according to whatever standard of goodness is adopted, given a certain condition an additional condition is better than what is conditionally obligatory, then on Lewis' analysis, whatever is conditionally obligatory given the original condition will be conditionally obligatory given the augmented condition. And yet surely augmented conditions which are better as well as those which are worse can override a conditional obligation.[9]

Lewis' proposal has also been criticized recently by Holly Goldman in "David Lewis' Semantics for Deontic

Logic"[10] and by Alan McMichael in "Too Much of a Good Thing: A Problem in Deontic Logic."[11] According to both Goldman and McMichael, Lewis' ranking of deontic alternatives with respect to goodness leads to difficulties. Both argue that Lewis' analysis is too strong. Conditional obligation sentences we believe to be true turn out to be false on his analysis. Let us evaluate these criticisms in turn.

Goldman criticizes a version of Lewis' theory of conditional obligation according to which

(76) p ⇒ q is true at world i ≡
"(A) there are no evaluable worlds in which p is true,
or (B) some p & q world is better from the standpoint of i than any p & not-q world."[12]

Her criticism is as follows:

> . . .application to actual cases shows that this definition fails to ascribe truth to statements of conditional obligation which are obviously true, because it fails to take adequate account of the affect of contingent features of the world on such obligations. Suppose I promise to return a borrowed book tomorrow. Clearly, if I do not return the book tomorrow, I ought to apologize. But now consider possible worlds in which I do not return the book tomorrow. It is better, other things being equal, not to break a promise. Thus the best worlds in which I fail to return the book tomorrow are surely worlds in which I do not thereby break a promise: either because

> I am released from that promise before tomorrow, or because I never made such a promise in the first place. But in either of these worlds, nothing is gained by my apologizing for not returning the book. So it appears that a world in which I do not return the book and do not apologize (e.g. a world in which I have been released from my promise) is at least as good as any world in which I do not return the book but do apologize. The statement "If I do not return the book tomorrow, I ought to apologize," turns out to be false according to definition [76]. Nevertheless this statement is true, because in the actual world I will not be released from the promise.[13]

The most notable feature of Goldman's counterexample is that the conditional obligation sentence she suggests,

(77) If I do not return the book tomorrow, then I ought to apologize

is only true in <u>this</u> world given that I have made the promise. She argues that (77) is false on Lewis' analysis since his analysis ignores contingent features of the world such as the facts that I will not be released from the promise and have made the promise. But analogously, her sentence (77) ignores those very same features and is true only when they are. The crucial point on Lewis' analysis is to restrict one's attention to deontic alternatives where the condition holds. Then we determine what would be best in those circumstances. Goldman's sentence (77), however, forces us to focus on the wrong set of deontic alternatives. It appears, then, that what goes wrong is not Lewis' theory but rather that the sentence expressing the conditional obligation misleadingly omits conditions which are clearly necessary for the truth of (77) and thus, justifiably, for the proper restriction of deontic alternatives in the semantic analysis.

In Goldman's example, if p stands for "I promise to return the book tomorrow" and r for "I do not return the book tomorrow" and q for "I ought to apologize," then we would say that p & r conditionally obligates one to q and $\sim$p & r does not. And these are just the results we would achieve on Lewis' theory. On his view the conditional obligation may change when the conditions change. His theory is carefully designed to accommodate that fact. But the result is, of course, that careful specification of conditions is required for accurate evaluation of conditional obligation sentences.

Goldman is aware of this sort of reply to her objection. She says,

> One might attempt to defend definition [76] by arguing that the conditional sentence "If I do not return the book tomorrow, then I ought to apologize," is merely a shorthand way of expressing what the speaker really means, which is more adequately expressed by the following statement: "If I have promised to return a book tomorrow, and I am not released from that promise, and I do not return the book, then I ought to apologize." The only version of the original statement which is completely immune to the sort of argument I have just advanced would require an antecedent describing <u>all</u>

> the features of the actual world which affect the suitability of my apologizing. But no speaker of ordinary English who utters the original sentence intends it to express a statement containing such an extended antecedent--an antecedent which may refer to indefinitely many facts, few of which would be known to him. Thus this defence of [76] fails.[14]

The defense raises the particularly difficult problem of distinguishing conditional and unconditional obligations. And Goldman's reply may be even more pressing than indicated here. For suppose we again attempt to defend Lewis as suggested above by arguing that Goldman's sentence (77) is an inadequate expression of the conditional obligation; the antecedent, or condition, needs to be supplemented in order to be a true conditional obligation sentence. But then, why not argue analogously that for any sentence of the form Op there are certain background conditions which must be true if Op is to be true, and these too must be incorporated into the obligation sentence if it is to be true? For example it may be patently false to claim it is obligatory not to kill, although if the killing is not in self defense it may be obligatory not to kill. If this argument is allowed then it appears that there can be no true unconditional obligation statements, but merely sentences of conditional obligation. Yet we have been maintaining that there is a distinction between conditional and unconditional obligations. Thus we discover that Goldman is correct; we can not require that conditional obligation sentences be supplemented if we do not wish the distinction between conditional and unconditional obligation sentences to collapse.

Interestingly, it seems clear to me that in Goldman's example the conditional obligation is conditional not only on not returning the book but also on promising to return it. The sentence makes no true statement of conditional obligation unless we take account of the assumption that I have promised either by adding it to the antecedent as a conjunct or by accounting for it in some other way. Similarly, Lewis' analysis requires that we incorporate the fact of the promise. And whether or not "all the features of the actual world which affect the suitability of my apologizing"[15] must be included in the antecedent, it seems that the fact that I have promised surely must be included.

But I see no way of justifying this view, nor do I see any way of distinguishing those features which must be included in the antecedent of the conditional obligation sentence. It is hardly fair to stipulate that we must include exactly those features which, when included in the analysis, yield intuitively acceptable results about the truth of the relevant conditional obligation sentence.

Thus I believe Goldman has singled out a serious weakness of Lewis' theory.

Historically, deontic logic, with its tripartite division between the obligatory, permissible, and forbidden, developed at about the same time as other logicians and ethicists developed formulations of act-utilitarianism separating duties, right acts, and wrong acts. The latter formulations generally distinguish a duty or obligation as the (possible) act of the available alternatives which produces (or would produce) maximal utility. The background conditions are built into the structure of specifying alternative acts.

For our project I have emphasized that I do not believe there is a way of specifying exactly which conditions must be included in the antecedent of a conditional obligation sentence. Nevertheless, a plausible way of keeping track of relevant (and then also irrelevant) background conditions on an obligation or conditional obligation is to build into the semantics certain restrictions narrowing accessible deontic alternatives. In particular, temporal restrictions may mark off many of the relevant conditions. For as time passes certain background conditions come to be true, and these should be held fixed in our theory. Goldman's sentence (77) is true at a certain time, namely at a time at which she has promised to return the book and has not been released from the promise. Thus incorporating into the semantic apparatus a temporal reference to reflect the fact that (77) is true at a particular time may help the semantic analysis more accurately evaluate obligation sentences. These comments provide only a rough introduction to the type of proposal we shall investigate in Chapter V.

It was clear from our discussion in Chapter I that statements about what agents conditionally or unconditionally ought to do can be expressed most clearly by requiring that such statements be relativized to both times and agents. Thus in developing an analysis for a concept of conditional ought-to-do, we must augment it with time and agent relativizations. And it may be that in so doing we gain the ability to make a clearer distinction between conditional and unconditional obligations. For what made it appear that the distinction between conditional and unconditional obligation sentences could not be made was the sweepingly general character of the latter. Nearly every example one can give of an unconditional or absolute obligation sentence seems to be implicitly conditional upon the situation. Yet if all obligation statements are relativized to agents and times then while we may not truthfully claim that it is unconditionally obligatory that one not kill, we may truthfully say that agent s (unconditionally) ought not kill at time t. For the references to agent and time may specify the situation carefully enough to assure us that the obligation statement has (as we have been assuming it has) a truth value. Exceptions due to extenuating circumstances will have no place. We shall explore this suggestion more carefully in Chapter V.

Another recent criticism of Lewis' analysis of conditional obligation has been presented by Alan McMichael. He also argues that Lewis' analysis is too strong; true conditional obligation sentences will be false on Lewis' account. He says,

> I am inclined to believe that there is a good, perhaps happiness is such a good, which may exist in amounts of any size. But if I am right, then all of the usual examples of conditional obligation fail to satisfy Lewis's truth-condition. Consider Lewis's example: given that Jesse robs the bank, it ought to be that he confesses and returns the loot (p. 102). Select any world _w_, however good, at which Jesse robs the bank (ϕ). Then there will be a world _w'_ which is better than _w_, but in which Jesse robs the bank without confessing and returning the loot ($\sim(\phi \supset \psi)$). This is so because in some such world _w'_, there will be enough extra good to counterbalance the absence of a confession and surpass the goods of _w_. In other words, since there is a good which may exist in amounts of any size, the lack of a confession puts no bound on the goodness of the worlds in which Jesse robs the bank. Consequently, Lewis's truth-condition yields the undesirable result that it is _not_ true that given that Jesse robs the bank it ought to be that he confesses and returns the loot.[16]

McMichael's point is that although Lewis believes his theory to be general insofar as it is suitable for any plausible ordering of the deontic alternatives, his theory is unacceptable for theories which are consistent with the view that there is a good which may exist in amounts of any size. Lewis has proposed a teleological analysis. Obligation and conditional obligation are defined in terms of the best worlds. The criterion for differentiating best worlds is an independent one. And this criterion may be one for which there are no limits to the values or goodness of best worlds.

Lewis is well aware that there may be, in this deontic context, an ascent to better and better worlds which is unlimited. (Cf. his discussion of the Limit Assumption, pp. 97-8 of Counterfactuals). Thus the criticism is not merely that worlds may be better without limit, but, more strongly, that even worlds in which p and ∿q hold, for any p and q, may get better without limit. Hence there is no point at which all the higher ranked worlds where p is true are worlds where p and q are true. Thus p never yields a conditional obligation to q for any p and q.

We might have expected Lewis to reply to McMichael that the better and better p and ∿q worlds are somehow excluded from consideration in the analysis by his evaluability restrictions. McMichael claims that references to standpoint and evaluability are dispensable if we are treating moral obligation. But perhaps the notion of a world being evaluable from another world could be made more specific. Perhaps a world w_2 could be said to be evaluable from the standpoint of a world w_1 just in case w_2 is "sufficiently similar" to w_1. Then worlds which differ from any other world only insofar as more good exists in them, would not be deemed "sufficiently similar." Lewis does not take this line, however.

Lewis responds to McMichael's objection[17] by granting that his semantic analysis does produce the results McMichael claims it produces when applied to a "strange doctrine" like the one McMichael suggests. But he maintains that this is the fault of the ethical doctrine appealed to, not his semantic analysis. He believes it is a virtue of his analysis that when applied to what he considers is a "strange" (counterintuitive?) doctrine, the theory produces "strange" (counterintuitive?) results.

The "strange doctrine" Lewis claims McMichael appeals to is characterized by Lewis as "radical utilitarianism, stark and unqualified," a view for which "agreement with our ordinary ethical thought is not its strong point." Yet surely this characterization of McMichael's theoretical point is oversimplified and thus misleading. Certainly one may be a dogmatic utilitarian and believe that what ought to be the case is whatever maximizes utility, and maintain further that however one defines utility, it is consistent with that definition that there be a good which exists in amounts of any size and which, when added to worlds, may raise their ranking in the "goodness ordering." And if the good in question is pleasure, this hedonistic assumption is a natural one for many utilitarians; indeed, it has been associated with utilitarians since Mill. But the assumption is consistent for non-hedonistic utilitarians as well. For a non-hedonist who believes that whatever ought to be is determined in terms of the good may define the good as love of beauty, truth, sympathetic impartial approval, knowledge, justice, love of virtue, or may even characterize it as a simple, non-natural quality (cf. Lewis' description of the goodness ranking for deontic alternatives, quoted in the beginning of this chapter). And for each of these is it not reasonable to grant that it may exist in amounts of any size? Even philosophers such as Brentano and Moore who denied that there is always a straightforward relationship between the value of a whole and the values of its parts, agree that there is a good which may exist in amounts of varying size.

When Lewis admits that deontic alternatives may be ranked according to some standard of good, then he is acknowledging that there is a standard of good which exists in unlimited amounts and McMichael's point follows. The argument is not restricted to "radical utilitarianism" or even hedonism.

Given the difficulties associated with Lewis' analysis of obligation in terms of best worlds, we shall begin a new approach to the problem of conditional obligation in Chapter V. Time shall again play a pivotal role in this approach.

NOTES TO CHAPTER IV

1. Bas van Fraassen, in "The Logic of Conditional Obligation," Journal of Philosophical Logic 1 (1972), pp. 417-438, proposes an analysis of conditional obligation which is similar to Lewis'. Van Fraassen, however, relies on a preference ordering of values ("items of unspecified character") that are realized at worlds rather than an ordering of worlds. Given van Fraassen's ordering of values at worlds, it is possible to derive an ordering of the worlds themselves (see Lewis, Counterfactuals, pp. 103-4). Lewis' analysis is equivalent to a special case of van Fraassen's (see Lewis, p. 104). Van Fraassen has since rejected his theory; cf. "Values and the Heart's Command," Journal of Philosophy 70 (January, 1973).

2. Lewis, David, Counterfactuals, p. 96.

3. Ibid., p. 97.

4. Ibid., p. 97.

5. Ibid., footnote, p. 100.

6. Ibid., p. 98.

7. Lewis, David, "Semantic Analyses for Dyadic Deontic Logic," p. 3.

8. Hansson, Bengt, op. cit., and see also David Lewis, "Semantic Analyses for Dyadic Deontic Logic," pp. 12-13.

9. This feature of Lewis' system has been noticed by others but I know of no other written articulation of the criticism.

10. Mind 86 (April, 1977), pp. 242-248.

11. Analysis 38 (March, 1978), pp. 83-84.

12. Goldman, op. cit., p. 247. This is equivalent to our version (68) above. For there is a world in the system of spheres such that p and q both hold there and such that any world as high or higher in the ranking in which p holds is a world where q holds if and only if there is a world where p and q hold and such that no world where p and ∿q hold is as high or higher

in the ranking than it is.

13. Goldman, _op. cit._, p. 247.

14. Goldman, _op. cit._, pp. 247-248.

15. _Ibid._, p. 247.

16. McMichael, Alan, "Too much of a good thing: a problem in deontic logic," _Analysis_ 38 (1978), pp. 83-84.

17. Lewis, David, "Reply to McMichael," _Analysis_ 38 (1978), pp. 85-86.

CHAPTER V

In Chapter I we developed a language for expressing time relativized obligation sentences by appealing to units of time which were assumed to be ordered linearly. That is, we assumed that given any two times t_1 and t_2, if $t_1 \neq t_2$ then $t_1 < t_2$ or $t_2 < t_1$. However for many philosophical purposes a non-linear temporal model may be more illuminating. In particular A.N. Prior first suggested, and Richmond Thomason has developed, an <u>indeterministic</u> or branching time structure as a model for tense logic.

> Such an account of time will permit instances in which a time α has alternative possible futures. The linear conception of time can countenance "alternative futures" as epistemic possibilities; although according to this conception each time α can have only one possible future, we may not have full knowledge at α of what the future will be like so that for all we know there are many alternative futures for α. These alternatives, however, being generated by incomplete knowledge rather than by the nature of time itself, do not enter into the truth-conditions for tenses. But nonlinear time puts these alternatives into the ontological structure of time, so that they must be taken into account in reckoning the truth-values of tensed formulas.[1]

The fundamental semantic idea behind this model theoretic technique is

> . . . that the formulas of a given formal language L will be assigned truth-values which may differ from time to time, and that the truth-values of tensed formulas at a given time will depend on the truth-values taken by that formula at other times.[2]

In two unpublished manuscripts, "Deontic Logic as Founded on Tense Logic" and "Deontic Logic and Freedom of Choice in Moral Deliberation," Thomason extends his indeterminist temporal model structure to provide a semantic analysis for certain formulas taken to express obligation statements. Thomason suggests that difficulties which arose in early deontic logic relevant to ethical questions such as questions relating to reparational oughts can only be solved with a combined tense and deontic language. He also claims in particular that his theory provides solutions to those difficulties.[3] Although Thomason does not go on to discuss this claim, it is an intriguing and challenging one. In this chapter we shall attempt to build a theory of conditional ought-to-do on his indeterminist time model.

Thomason first distinguishes between two uses of "ought." The <u>deliberative</u> use is appropriate for advising or deliberating while the <u>judgmental</u> ought is appropriate for passing judgment or wishing. This distinction can be illustrated with an example.

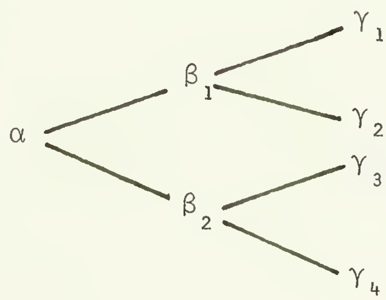
> Say that I borrow from you some notes that you need for a lecture you're giving at 2:00, and promise to give them back in time for your lecture. I take them to my home, half an hour away from where the lecture is to occur, and sit down to spend time. Up to 1:30 I ought (on the deliberative reading) to return the notes by 2:00, and we can even measure the increasing strength and urgency of the ought as time wears on toward 1:30, by the time available before it will become impossible for me to fulfill the ought . . . At 1:30 or thereabouts, this deliberative ought is replaced by a judgmental ought (good only for condemnation, since by this time it's impossible for me to return the notes before the lecture), and also gives way to a deliberative reparational ought of some sort--perhaps I ought to call you and provide some warning, or perhaps I ought to take the notes directly to the lecture. As time goes on, if I ignore them these reparational oughts will in turn be replaced by further condemnations.[4]

The ought of judgment thus arises from a wider set of alternatives than the ought of deliberation. To a recording angel who is judging me with respect to the alternatives that have been open to me at 1:45 it is clear that I ought to return the notes at 2:00. But the deliberative ought is determined from a narrower set of alternatives, namely those actually open to me or in my power at the time. Thus at 1:45 it is false to say I (deliberatively) ought to return the notes. A deliberative ought must be possible for me in some sense of "possible."

The terms "deliberative" and "judgmental" suggest that Thomason's distinction might be very like the traditional distinction between what ought to be and what an agent ought to do.[5] However given Thomason's illustration it seems more likely that both uses of "ought" are uses of the ought-to-do. Perhaps the deliberative use is appropriate for situations in which one is choosing what to do from various alternatives that are in fact open at the time while the judgmental use is suitable for evaluation of what ought to have been done at instants in the past. If so, then it would seem that an analysis of the judgmental use of "ought" could be developed in terms of the deliberative use of "ought." Indeed, Thomason suggests that this is so.[6] Whether or not this development can be carried out successfully is an open question at this point. Still, if this rough characterization of Thomason's uses of "ought" is accurate and time provides a way of drawing the distinction, then clearly we must focus on the deliberative use of "ought" to develop meaningful prescriptions about how an agent ought to act.

Our concern, then, is with the deliberative use of "ought" for which certain alternatives which might have been open to me are no longer available. Some are ruled out by physical circumstances, perhaps others by psychological factors. But the main point is that many are ruled out as unavailable options as time passes. Hence we may begin to see the plausibility of building a theory of conditional ought-to-do on the indeterministic time model, according to which available alternatives may vary from time to time.

Thomason's formal semantics is based on a theory in which time may branch toward the future. There may be two times, both in the future with respect to a given time, that are temporally unrelated to each other. These two times, then, are on different "branches" of the time structure. If we assume that time may be partitioned into discrete units we may have treelike structures as in the following finite illustration:



At α there are four possible future courses of events or <u>scenarios</u>. At β_1 and β_2 there are two future scenarios. We may designate the instants by Greek letters to emphasize that they are non-linear units of time. According to Thomason "the fact that instants β_1 and β_2 are incomparable signifies that realization of β_1 will exclude those alternatives in which β_2 is realized."[7] Presumably if β_1 and β_2 are incomparable then it will not be the case that $\beta_1 = \beta_2$ or that $\beta_1 < \beta_2$ or that $\beta_2 < \beta_1$. This theory of time "allows certain future alternatives that formerly were open to become extinguished with the flow of time,"[8] which is just what we require for an analysis of the ought of deliberation.

Thomason begins with a nonempty set K of (non-linear) discrete times, or "reference points" at which formulas will be evaluated, ordered by a relation < (to be interpreted as an "earlier than" relation.) Then a _temporal model structure_ is defined as a pair $\langle K, < \rangle$, where the most important condition of the structure is that for all $\alpha, \beta, \gamma \in K$, if $\beta < \gamma$ and $\alpha < \gamma$, then $\beta < \gamma$ or $\gamma < \beta$ or $\beta = \gamma$. This condition assures us that time can not branch into the past. Moreover, < is transitive, that is, if $\alpha < \beta$ and $\beta < \gamma$, then $\alpha < \gamma$. Finally, Thomason requires that for all $\alpha \in K$, there is a $\beta \in K$ such that $\alpha < \beta$.[9]

A _history_ h on a model structure is a subset of K such that

1) for all $\alpha, \beta \in h$, if $\alpha \neq \beta$ then $\alpha < \beta$ or $\beta < \alpha$, and

2) if g is any subset of K such that for all $\alpha, \beta \in g$, then $\alpha < \beta$ or $\beta < \alpha$, then $g = h$ if $h \subseteq g$.

A history is a maximal chain on a structure, that is, a linear pathway through the structure that skips no instants. If $\alpha \varepsilon K$ then H_α is the set of all histories containing α. If $h \varepsilon H_\alpha$, the segment of h beyond α corresponds to a possible future for α. Since time branches only toward the future, then, H_α represents the set of scenarios open at α.

We may focus on a subset of Thomason's formal language L including atomic sentence variables P,Q,R,. . ., and sentence connectives $\sim$, $\supset$, F (for future tense) and O (for obligation). A valuation V assigns truth-values $V_\alpha^h(A)$ to formulas A of the given language L at an instant α relative to a history h in H_α. $V_\alpha^h(A)$ defines the truth-value taken by A at α, <u>provided</u> that history h includes what will happen after α. Thus the valuation function gives the truth-value of A at a time assuming that a certain course of events will come to pass. This valuation function is defined by induction on the complexity of A. For an atomic sentence letter P,

$$(78) \quad V_\alpha^h(P) = V_\alpha(P),$$

and, for example, for a formula A,

$$(79) \quad V_\alpha^h(\sim A) = T \text{ iff } V_\alpha^h(A) = F.$$

Thomason captures Prior's "Ockhamist" theory of the future tense by defining

(80) $V^h_\alpha(FA) = T$ iff $V^h_\beta(A) = T$ for some $\beta \in h$ such that $\alpha < \beta$.

If there is more than one course of events h in H_α, $V^h_\alpha(A)$ will not reflect the unqualified truth-value $V_\alpha(A)$ taken by A at α. Thus Thomason defines $V_\alpha(A)$ using van Fraassen's method of truth-value gaps:

(81) $V_\alpha(A) = T$ iff $V^h_\alpha(A) = T$ for all $h \in H_\alpha$,

(82) $V_\alpha(A) = F$ iff $V^h_\alpha(A) = F$ for all $h \in H_\alpha$

and

(83) $V_\alpha(A)$ is undefined otherwise.

These truth definitions which allow for truth value gaps are especially useful for the semantic analysis of future tense statements. They provide a formal development of the view that "future contingent" statements are neither true nor false.

In order to provide a semantic analysis for the deontic operator O, interpreted deliberatively, Thomason introduces a model structure $\langle K, <, S \rangle$ where S is a relation between instants and histories such that if αSh then $h \varepsilon H_\alpha$. The semantic determinant S is used to interpret the deontic operator O. Thomason says, "intuitively, the meaning of αSh is that h represents a future course of events at instant α that would be an acceptable moral choice."[10] Those histories related by S to an instant α form the ought set of α. Thomason's semantic rule for the deontic operator, interpreted deliberatively, is

$$(84) \quad V_\alpha^h(OA) = T \text{ iff } V_\alpha^g(A) = T \text{ for all } g \text{ such that } \alpha Sg,$$

or equivalently,

$$V_\alpha^h(OA) = T \text{ iff } (g)(\alpha Sg \supset V_\alpha^g(A) = T),$$

with $V_\alpha(OA)$ defined using (81) - (83).

This analysis gives O an S_5-like structure.[11] The truth-value of a formula OA with respect to a history h and a time α depends solely upon the truth-value of A at α in those histories in the ought set of α. The history of evaluation is therefore irrelevant for the deontic operator. That is, if $V^{g}_{\alpha}(A) = T$ for all g such that αSg then $V^{h}_{\alpha}(OA) = T$ for all h ε H_{α} and hence $V_{\alpha}(OA) = T$. When it is not the case that $V^{g}_{\alpha}(A) = T$ for all g such that αSg then $V^{h}_{\alpha}(OA) = F$ for all h ε H_{α} and hence $V_{\alpha}(OA) = F$. Obligation statements will thus be either true or false. When we supervaluate according to (81) - (83), a formula OA will never lack truth-value. Thus for any history h, $V^{h}_{\alpha}(OA) = T$ iff $V_{\alpha}(OA) = T$, and so (84) reduces to

(85) $V_{\alpha}(OA) = T \text{ iff } (g)(\alpha Sg \supset V^{g}_{\alpha}(A) = T)$.

The idea is apparently that OA is true at α in every alternative future which represents a morally acceptable course of events from the point of view of α. That is, A would be true in all morally acceptable histories containing α if those histories were actual, regardless of which course of events does follow.

But upon reflection this explanation (and thus (84) and (85)) does not appear to be as clear as it may at first have seemed. For in evaluating a formula OA at a time α we merely check the truth-value of A at α. Thus it is difficult to see how any future alternatives, even those which are morally acceptable, are relevant to the analysis.

Indeed, for a sentence variable P, the truth of OP at α depends solely upon the truth-value of P at α. Similarly, if A is not atomic, but contains no future tense operators, the evaluation of OA at α still depends only on the truth-value of A at that same time. There is no reference to a different time and thus again reference to future alternatives is irrelevant. Thus if A contains no future tense operators, then

(86) $(\alpha)(V_{\alpha}(OA \equiv A) = T)$.[12]

We do not wish to affirm in general that whatever is obligatory is true.[13] Hence a formalization of straightforward obligation statements may not be symbolized using OA, where A is atomic or has no future tense operators, in Thomason's system, contrary to what we may have believed and what Thomason suggests in some examples.[14] Rather, Thomason's formal expression of the English sentence "I ought to pay," for example, must employ a future tense operator. If P stands for "I pay," he must write as the formalization, OFP. Then "I ought to pay" is true at a time α just in case every morally acceptable history through α is such that there is a time on it after α when I pay. Using this formalization we may truly say that I ought to pay is true at α even though it is not true that

I pay at α, which is as we wish.

This does seem to be Thomason's intention and perhaps it shows one way of understanding his claim that deontic logic is intimately tied to and reliant upon tense logic.[15] The interesting obligation sentences in his system are of the form OFA. And OFA may be true when FA lacks truth-value.

However we still seem to have a difficulty. For Thomason uses as examples of English sentences which express obligations that we wish to formalize,

(87) George provides transportation Saturday,

and

(88) I ought to give you $50 tomorrow.

Thomason says,

> To take an example, let _A_ correspond to 'I will give you $50 tomorrow' and suppose that I owe you the money, have promised to give it to you, and can do so--there are scenarios ahead of me on which I pay you the money tomorrow. Then I would want to say that O_A_ is true--I ought to give you the money tomorrow--and so every scenario in my ought set is such that on that scenario I _do_ give you the money tomorrow.[16]

Although indexicals such as "tomorrow" or "now" pose special problems, it does seem that we would like to symbolize sentences such as

(89) I ought to pay you Saturday at 3:00

(where we even include a calendar date if the reference to Saturday is ambiguous. Let us assume for our examples that we need not add reference to a calendar date.) And there does not appear to be an adequate symbolization for (89) available in Thomason's system.

Imagine that (89) is true at α. If we first attempt to let the atomic sentence letter P stand for "I pay you Saturday at 3:00" and use OP as the symbolization for (89), as seems to be suggested by Thomason in his example above, then (86) holds in this case. With this symbolization it follows from (89) that it is true at α that I pay you Saturday at 3:00, which, given my fallibility, may clearly be false.

It is natural, then, to attempt to symbolize (89) by using the future tense operator F so that (86) will not hold.

(90) FOP

is an implausible candidate. For we wish to assert the truth of (89) at α and so wish to consider the ought set or morally acceptable histories open <u>at α</u>, whereas (90) is true at α just in case

$$(91) \quad (h)(h \in H_\alpha \supset (\exists\beta)(\beta \in h \,\&\, \alpha < \beta \,\&$$

$$(g)(\beta S g \supset V^g_\beta(P) = T))),$$

that is, just in case on every history through α there is a time later than α in which the ought set has certain features. Hence we might try to symbolize (89) using

(92) OFP

where P stands for either "I pay you" or "I pay you Saturday at 3:00."

Consider first that P corresponds to "I pay you." Then if (92) symbolizes "I ought to pay you" as we have argued above, it can not be adequate as a formalization for (89). These two English sentences are not equivalent and so identical symbolizations can not be satisfactory. The difficulty seems to be that even though (92) indicates that in each history which is morally acceptable at α there is a later time at which I pay you, nothing in symbolization (92) captures what is expressed in the English sentence, namely that it is on Saturday at 3:00 that I ought to fulfill the obligation.

Suppose then that P stands for "I pay you Saturday at 3:00." Then (92) is true at α just in case

(93) $(g)(\alpha Sg \supset (\exists\beta)(\beta \varepsilon g \ \& \ \alpha < \beta \ \& \ V_{\beta}^{g}(P) = T))$,

that is, just in case all histories which are morally acceptable at α are such that there is an instant along them later than α at which it is true that I pay you Saturday at 3:00.

This suggested symbolization (92) is closer to what we would like than the previous ones and also seems in line with Thomason's view. However we may still ask how we are to evaluate $V_{\beta}^{g}(P) = T$ if P stands for "I pay you Saturday at 3:00." Thomason addresses this difficulty in a footnote:

> I want to think of (2.4) ['George provides transportation on Saturday'] as true at any time if and only if 'George provides transportation' is true on Saturday.[17]

Analogously, for a given history, "I pay you Saturday at 3:00" is true at any time in the history just in case "I pay you" is true at 3:00 on Saturday in that history. Notably, we can not express this in Thomason's language. For P symbolizes "I pay you Saturday at 3:00" but we have no symbol for "I pay you." More importantly, we have no way of referring to Saturday at 3:00. It is what we might call (recalling a phrase of Thomason's) a "clock-time" and is not a time in the sense that α, β, and γ . . . are times. For each of α, β, and γ . . . is an instant occurring in at most one position in the branching time structure. Yet presumably there is a time corresponding to Saturday at 3:00 in every history open at α.

Unfortunately there is no way of formulating in Thomason's language the general principle which he advocates in the footnote. If we attempt to formulate it by reverting to let P stand for "I pay you" and P_3 for "I pay you Saturday at 3:00" then it appears that he is saying

(94) $(\alpha)(V^h_\alpha(P_3) = T \text{ iff } V^h_3(P) = T)$.

But (94) is not well formed in his language. As pointed out above, Saturday at 3:00 is not a time in the sense that α, β, and γ . . . are times; it can not serve as a reference for the valuation of P.

Despite the inadequacy of his language for expressing the principle, the intuition behind it is a plausible one, and is commonly appealed to in explanations of the truth of such time-relativized sentences. However, if we accept

(92) OFP

where P corresponds to "I pay you Saturday at 3:00" as a symbolization for

(89) I ought to pay you Saturday at 3:00,

and accept the intuition underlying Thomason's informally stated principle, then we are forced to an unacceptable consequence.

Whenever V^h(OFP) is true then for every g such that αSg there is a β, $\beta \varepsilon$ g, $\alpha < \beta$, such that $V^g_\beta(P) = T$, (see (93)). Since αSg we know that g ε H_α. Thus P must have a truth-value at α in g. According to Thomason's principle, applied with respect to a history, if P ("I pay you Saturday at 3:00") is true at a time in that history, then P is true at any time in that history. Hence, if $V^h_\alpha(OFP) = T$, then if we let g_1 be a history such that αSg_1, there is a $\beta_1 \varepsilon g_1$ such that $V^{g_1}_{\beta_1}(P) = T$. And so by Thomason's principle as stated above, $V^{g_1}_\alpha(P) = T$. But notably P is an atomic sentence variable. And so according to (78), the truth-value of P at a time relative to a history is just its value at that time. Thus we have $V_\alpha(P) = T$.

In sum, if $V^h_\alpha(OFP) = T$, and hence $V_\alpha(OFP) = T$, then if

(92) OFP

symbolizes

(89) I ought to pay you Saturday at 3:00

where P stands for "I pay you Saturday at 3:00," it follows that by appealing to Thomason's informally stated principle we can derive $V_\alpha(P) = T$. If "I ought to pay you Saturday at 3:00" is true at α then "I pay you Saturday at 3:00" is true at α. This is clearly an unacceptable consequence.

One might suggest that Thomason should not have used as examples sentences like (87) and (88). For if we restrict our attention to sentences of the form "I ought to pay you," Thomason's proposal is immune to the difficulty above. His theory is unproblematic as long as we are careful to understand that the atomic sentence variables P, Q, R, . . . can not symbolize English sentences which contain references to particular clock-times such as Saturday at 3:00. However I believe we should have a symbolization and semantic analysis for sentences like (89) and Thomason's examples indicate that he believes so as well. One might also suggest that Thomason drop the principle he has stated informally in a footnote. However if we are to symbolize sentences of the form of (89) then some version of the principle is crucial and I have claimed it is plausible as well. Yet I see no way of formulating it plausibly in Thomason's system.

There may be a way of preserving Thomason's system and (*1*) symbolizing sentences of the form of (89), (*2*) maintaining and formalizing the intuitive idea underlying his informally stated principle, and (*3*) avoiding the difficulty presented above, and I explore that possibility in Appendix II. But for the remainder of this chapter we shall be careful to recall that we must restrict the English sentences which may be symbolized by atomic sentence variables (thereby restricting the obligation sentences of English that we may express) in Thomason's language by excluding those sentences which contain a reference to a particular date and time.

We have made no attempt thus far to provide any agential references in our semantic analysis for obligation sentences. Thomason presents (84) as the semantic rule for the deliberative use of "ought," but he has not incorporated any reference to the agent for which the obligation may be said to hold. Rather, he has presented a one-person theory; that is, he has assumed that the obligation statements hold for a single agent. However Thomason is well aware that different agents have different obligations in the deliberative sense.[18] Thus it will not do to refer to a general, impersonal ought set at a time.

Rather, at any given instant, every individual has his or her own ought set. Thomason makes this suggestion in the illustration we gave on page 138. He says there that he gives the \$50 in every scenario in his ought set. Thus it is his ought set that is used to determine what he ought to do. At another point he claims, more specifically, that "everyone should have his or her own ought set."[19]

Thus we might amend Thomason's (84) to state that OA is true at a time in a history for an agent just in case A is true at that time in all scenarios which are acceptable moral choices for that agent (that is, those scenarios in the agent's ought set) at that time. Thus if s is a variable ranging over agents named by s',s", . . ., and S_s is a relation between instants and histories such that if $\alpha S_s g$ then $g \in H_\alpha$, and, intuitively, g is in the ought set for s at α, then we might state (84) as

$$(95)\ V^h_{s,\alpha}(OA) = T \text{ iff } (g)(\alpha S_s g \supset V^g_\alpha(A) = T).$$

However, valuations of non-deontic sentences of the language are determined solely with respect to times and histories. Ultimately, using (81) - (83), every formula is evaluated solely with respect to a time. To maintain this feature for obligation sentences it is preferable to define truth-values for sentences, not of the form OA, but of the form O_sA, which may be read "A is (deliberatively) obligatory for s" or "s ought to see A."[20] Then we shall have

$$(96)\quad V_{\alpha}^{h}(O_sA) = T \text{ iff } (g)(\alpha S_s g \supset V_{\alpha}^{g}(A) = T),$$

where h is irrelevant.

Let us now attempt to develop an analogous semantic analysis for conditional obligation as an ought-to-do and deliberative ought, and let us attempt to formulate it for sentences of the form "Smith is conditionally obligated to repent given his sin," where no mention is made to specific clock-times.

Thomason's truth definition for unconditional obligation sentences tells us that A is obligatory for s at α just in case A is true in all those histories which are morally acceptable for s at α. Thus, recalling other analyses of conditional obligation that we have discussed, we might wish to say that C is conditionally obligatory given B for s at α just in case C holds in all those histories which are morally acceptable for s at α and in which condition B holds.

This proposal will not do, however. At any given time α there is, on Thomason's view, a set of histories which are morally acceptable for s at α, namely s's ought set at α. Thus, for example, if it is true at α that s ought not to sin, it is true that s does not sin in all those histories in his ought set. But then there can not be a subset of those histories in which he sins. Hence it is vacuously true that any C is conditionally obligatory for s given that s sins. On this proposal the forbidden conditionally obligates one to anything.

We might attempt to block vacuous truth of conditional obligation sentences by altering the above proposal to say that C is conditionally obligatory given B for s at α just in case (*i*) C holds in all those histories which are morally acceptable for s at α and in which condition B holds, and (*ii*) there is at least one history morally acceptable for s at α in which condition B holds. But this proposal is also flagrantly inadequate. For as in the case above, there will be no acceptable histories for s at α in which s sins (assuming s ought not to sin at α) and so we can not express a true conditional obligation at α for s to repent for a sin. No contrary-to-duty obligations will be true on this analysis since in those cases condition (*ii*) can not be satisfied.

To avoid these unwelcome results let us try another approach. We might rather sloppily say that C is conditionally obligatory given B for agent s at time α just in case C is true in every α-history in which there is a time such that condition B is true then and which is in the agent's ought set with respect to that time.

To clarify this intuitive idea, let us consider an example. Let us suppose Smith now ought not to sin. His ought set now contains no future scenarios in which he sins. But we wish to maintain now that he ought to repent if he does sin. Thus it seems reasonable to claim that once Smith does sin he repents in all the histories in his ought set at the time of his sinning. After his sin those histories in which he does not sin are no longer available; they have been bypassed. The morally acceptable futures available to him at the time of his sinning are such that he repents in them. A proposal based on this intuitive idea is clearly more plausible than our first proposals and will capitalize on the feature of the branching time model that available alternatives may vary from time to time.

More specifically, if s is a variable ranging over agents, and α and β range over times or instants, and f, g, and h are variables ranging over histories, then for formulas B and C we might suggest:[21]

(97) $V^h_\alpha(O_s(B/C)) = T$ iff

$$(g)(\beta)\{(g \in H_\alpha \ \& \ g \in H_\beta \ \& \ V_\beta^g(B) = T) \supset$$

$$(f)(\beta S_s f \supset V_\beta^f(C) = T\},$$

which is equivalent to

$$V_\alpha^h(O_s(B/C)) = T \text{ iff}$$

$$(g)(\beta)\{(g \in H_\alpha \ \& \ g \in H_\beta \ \& \ V_\beta^g(B) = T) \supset$$

$$V_\beta^g(OC) = T)\}.$$

On this proposal, like Thomason's (84), conditional obligation statements will not lack truth value when we supervaluate using (81) - (83). Hence h is irrelevant again.

Unfortunately, if this proposal is intended recursively, and is thus intended to apply to any formulas B and C, it is unsuitable for at least two reasons. As before, the interesting cases will be those for which C and perhaps B as well, are future tense formulas which lack truth-value. And yet if we try to symbolize Chisholm's sentence that it is now true that if a certain man goes to the assistance of his neighbors he is conditionally obligated to tell them he is coming, where s' names the man, and G and

T abbreviate "s' goes to assist his neighbors" and "s' tells them he is coming," using

(98) $O_{s'}(FG/FT)$

we can see two major problems.

First, the analysis tells us that at every time and history at which the condition is true the acceptable moral histories at those times are such that FT is true then. Thus all histories morally acceptable at a time when the condition holds are such that there is a later time when s' tells them he is coming. And yet clearly we would like our analysis to allow us to assert that Chisholm's sentence is true and that the conditional obligation, telling, may occur before the condition, going. Given (97) we can not symbolize true conditional obligation sentences for which the conditional obligation is to occur before the condition.[22]

Second, the analysis in (97) directs us to every history g and time β such that $g \in H_\alpha$ and $g \in H_\beta$ and $V^g_\beta(FG) = T$. But if FG is true at any particular time β in g, because G is true at a particular time γ later than β, then surely FG is true at many other times in history g. We wish to look at the histories which are morally acceptable at those times when he goes, that is, when G is true, but not at all those times when it is true that he will go, that is, when FG is true. Using (97) we are forced to focus on too many times at which a given formula expressing the condition is true.

It appears to me, then, that the general truth definition (97) can not capture the intuitive idea we are trying to express about conditional obligation sentences. However, even if (97) fails as a general definition, we might propose that for atomic sentence variables P and Q,

(99) $V^h_\alpha(O_s(P/Q)) = T$ iff $V^h_\alpha(P) = T \supset$

$(f)(\alpha S_s F \supset V^f_\alpha(Q) = T)$.

And in the more interesting cases,

(100) $V^h_\alpha(O_s(P/FQ)) = T$ iff $V^h_\alpha(P) = T \supset$

$(f)[\alpha S_s f \supset (\exists\gamma)(\alpha < \gamma \ \& \ f \in H_\gamma \ \&$

$V^f_\gamma(Q) = T)]$

and

(101) $V_{\alpha}^{h}(O_s(FP/FQ)) = T$ iff

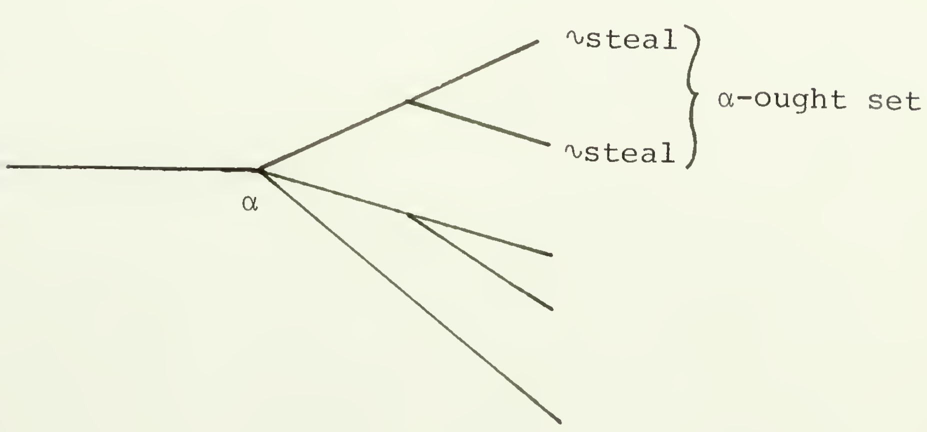
$(g)(\beta)\{(g \ \varepsilon \ H_{\alpha} \ \& \ g \ \varepsilon \ H_{\beta} \ \& \ \alpha < \beta \ \&$

$V_{\beta}^{g}(P) = T) \supset$

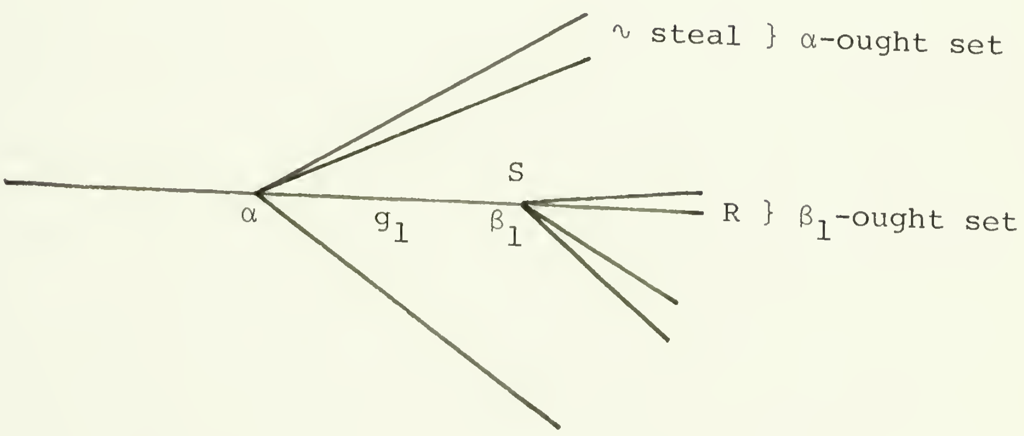
$(f)[\beta S_s f \supset (\exists\gamma)(\alpha < \gamma \ \& \ f \ \varepsilon \ H_{\gamma} \ \&$

$V_{\gamma}^{f}(Q) = T)]\}.$

Some diagrams may help us illustrate (101). Suppose that at α Smith has an obligation to return $100 if he should steal it. Surely his ought set at α contains no histories in which he steals the money since he ought not to steal it.



However, if S corresponds to "Smith steals \$100" and R to "Smith returns \$100," then consider a particular history g_1 through α and a time β_1 on that history such that $V^{g_1}_{\beta_1}(S) = T$. Then every history f in Smith's ought set with respect to β_1 is such that there is some instant γ such that $V^{f}_{\gamma}(R) = T$.



Again, in the final valuation the history of evaluation, h, is irrelevant. Moreover, since we are proposing an analysis of conditional ought-to-do, it is in line with our earlier discussions (Chapter I, p. 21 and Appendix II,

p. 204) to require that $\alpha < \gamma$ in (100) and (101).

If the conditional obligation in question is to repay stolen money or to repent for a sin then presumably the returning comes after the stealing and the repenting comes after the sinning, and so we have $\beta < \gamma$. However if we do not require this relationship between β and γ, then our analysis is applicable to cases like Chisholm's in which the conditional obligation for the agent to tell his neighbors he is coming should be fulfilled before he goes to help them. Thus we appear to avoid the first difficulty discussed with respect to (97).

Furthermore, if we restrict the sentences symbolized by the atomic sentence variables P, Q, R . . . to those English sentences which are such that if ever true in a history, they are true only once in that history, then these proposals avoid the second difficulty associated with (97).[23]

But (99) - (100) allow statements of conditional obligation to be true vacuously. For example, for the most interesting case, (101), if there is no history g and time β such that $g \in H_\alpha$ & $g \in H_\beta$ & $\alpha < \beta$ & $V^g_\beta(P) = T$, then O(FP/FQ) will be vacuously true for any Q. Thus an impossible condition, a condition which is not true in any possible alternative future, yields a conditional obligation to anything. Hence we might amend (99) - (101)

as we amended Lewis' proposal (from (70) to (71), pp. 102-103) by adding an existential clause ensuring that there is an alternative future in which the condition and what is conditionally obligatory do hold.

For example, for (101) we might now say

(102) $V^h_\alpha(O_s(FP/FQ)) = T$ iff

i) $(g)(\beta)\{(g \in H_\alpha \ \& \ g \in H_\beta \ \& \ \alpha < \beta$

$\& \ V^g_\beta(P) = T) \supset$

$(f)[\beta S_s f \supset (\exists\gamma)(\alpha < \gamma \ \& \ f \in H_\gamma$

$\& \ V^f_\gamma(Q) = T)$

& ii) $(\exists i)(\exists\sigma)(\exists\eta)(i \in H_\alpha \ \& \ i \in H_\sigma$

$\& \ i \in H_\eta \ \& \ \alpha < \sigma \ \& \ \alpha < \eta \ \& \ V^i_\sigma(P) = T$

$\& \ V^i_\eta(Q) = T)$.

Given (102), if there is <u>no</u> history along which P holds at a time later than α and along which Q also holds at some time after α, then we get the appropriate result that it is not true to say $O_s(FP/FQ)$ is true for any agent s.

One consequence of Thomason's truth definition for obligation sentences is that whenever A is true, OA is also true. One way of seeing why this is so is by introducing an operator L for inevitability into the language. The semantic rule for this operator is

$$(103) \quad V_\alpha^h(LA) = T \text{ iff } (g)(g \in H_\alpha \supset V_\alpha^g(A) = T).$$

A sentence LA is true at α if A's truth at α is independent of scenarios at α. Then, as Thomason points out, A semantically implies LA. Whatever is presently true is presently inevitable. Furthermore, in deliberative contexts, LA implies OA, for if A is true in all possible futures open at the time, A is clearly true in that subset of alternative histories available at the time which are morally acceptable. Thus, by transitivity, A implies OA. Whatever is true is obligatory.[24] Similarly, given the proposals we are discussing, if whatever is conditionally obligatory is true, it is conditionally obligatory given any possible condition.

When O is construed deliberatively this may seem somewhat peculiar, for it is difficult to imagine advising someone to see to something which is already the case. However these odd results merely indicate that on these analyses past and inevitable actions are classified as morally acceptable in the deliberative sense. We might not judge that all our past actions have coincided with what a recording angel would deem to be obligatory for us, but it is reasonable to discover that such a view must be expressed using the judgmental use of "ought." Furthermore, in deliberative contexts, obligations are determined by choosing among possible futures open at the time of deliberation. Thus if LA is true, it is not open to the agent to choose $\sim$A, and so we would not wish to say $\sim$A is obligatory (or conditionally obligatory on any condition) in the deliberative sense.

There are other interesting consequences of the proposals given above. If for atomic sentence letters P and Q, $V_\alpha(O_s(P/Q)) = T$ and the condition is true at α, that is, $V_\alpha(P) = T$, then it follows from (81) and (99) by modus ponens that

$$\text{(104)} \quad (f)(\alpha S_s f \supset V^f_\alpha(Q) = T),$$

and so that

$$\text{(105)} \quad V_\alpha(O_s Q) = T.$$

Similarly, if $V_\alpha(O_s(P/FQ)) = T$ and $V_\alpha(P) = T$ then applying (81) and (100) and modus ponens it follows that

(106) $(f)(\alpha S_s f \supset (\exists\gamma)(\alpha < \gamma \;\&\; f \in H_\gamma \;\&$

$V^f_\gamma(Q) = T)),$

from which we can conclude

(107) $V_\alpha(O_s FQ) = T.$

Hence we have

(108) $V_\alpha(O_s(P/Q)) = T \;\&\; V_\alpha(P) = T \supset$

$V_\alpha(O_s Q) = T$

and

(109) $V_\alpha(O_s(P/FQ)) = T \;\&\; V_\alpha(P) = T \supset$

$V_\alpha(O_s FQ) = T.$

If $V_\alpha(O_s(FP/FQ)) = T$ and $V_\alpha(FP) = T$, then we may not derive (107), an obligation <u>at α</u> for s to see to Q in the future. However it will follow that if FQ is obligatory for s given FP and FP is true, that is, P is true in every future history, then every history is such that there is a time on it at which Q is true (earlier or later) in every member of s's ought set.[25] If the condition holds sooner or later in every future, then sooner or later in every future there is a time at which whatever is conditionally obligatory is true at some time in each scenario in the agent's ought set. This seems to me to be a welcome consequence.

These results, indicating the relationships between certain conditional and unconditional obligations, are not susceptible to the difficulties, pointed out by Greenspan (see Chapter I, pp. 3 ff.), which arise for a general factual detachment rule. To return to her example, given a general factual detachment rule, the truth of the condition that I will get a ticket by the end of next month is sufficient to allow detachment of an obligation to pay the fine, and this is unacceptable.

Clearly, if P symbolizes "I get a ticket" and Q symbolizes "I pay the fine" then if I get the ticket at α, the time at which the conditional obligation holds, then from (109) we can detach an obligation at α to pay the fine (some time in the future).

Also, if FP symbolizes "I will get a ticket" and this is a true future tense statement, then in every future history I do get a ticket; getting a ticket is inevitable. In such a case it will follow that in every history there is a time when my ought set is such that in each of its members I do pay the fine. This is reasonable in deliberative contexts since it is not possible for me not to get a ticket.

If, on the other hand, FP symbolizes "I will get a ticket" and this is a future contingent statement, then V_α(FP) will be undefined and nothing whatsoever will follow about whether or not I eventually pay the fine in alternative futures.[26]

Our discussion of proposals (99) - (102) illustrates many of the advantages of building a theory of conditional obligation within Thomason's indeterminist time model. Conditional and unconditional obligations may be relativized to times, clarifying what is settled when an obligation is in force. If I have already made a promise, then that fact will be settled; any subsequent obligations or conditional obligations will arise from future options open to me, and not promising is not one of those. Moreover, since future alternatives open to me may vary from time to time, obligations and conditional obligations may also vary over time. This feature of the model makes it particularly appropriate for expressions of contrary-to-duty obligations such as our paradigm, an obligation to repent for a sin. Finally, our discussion has shown that analyses built on this model, unlike most dyadic analyses of conditional obligation, may allow for acceptable factual detachments of unconditional obligations from conditional ones, without allowing those detachments which cause difficulty for general factual detachment in standard deontic logic.[27]

But there are serious difficulties with these proposals. Recognition of these problems may provide some insight into a more appropriate way of developing a theory of conditional obligation within the indeterminist time model.

Due to difficulties with (97) we proposed (99) - (101) as non-recursive definitions for atomic and future tense sentences. And while many paradigm conditional obligation sentences can be analyzed appropriately using (100) and (101) (amended with an existential clause as in (102)), these limited semantic rules are clearly insufficient.

In particular, it might seem that a further advantage of the intuitive idea underlying these proposals is that even if Q is true in every history morally acceptable for the agent at the times when P holds, Q may not be true in every morally acceptable history open at those times when

P & R hold. Thus, to return to our example of Chapter I, it seems that we can say, as Greenspan can not, that if Smith charges a shirt at Macy's, it is conditionally obligatory for him to pay for it, although given that he charges it and returns it he may no longer be obligated to pay for it. Even if he pays for it in every history morally acceptable at the times when he charges it, he may not pay for it in acceptable futures at those later times when he charges and returns the shirt.

However to substantiate this claim we must formulate an analysis for conditional obligation statements with conjunctive conditions. To cover the interesting cases we shall also need to include a future tense operator. And if the analysis is to be analogous to (99) - (101), we must agree that for any P and R, if P and R symbolize sentences true at most once in a history, then P & R symbolizes a sentence true at most once in a history.

Yet I see no way of providing a rule analogous to (99) - (101) for conditional obligation sentences with conjunctive conditions which gives us the results we wish. For we have agreed to let P correspond to a sentence which is true at most once in a history. Yet this assumption prevents us from allowing conditional obligations to vary when a condition is augmented. This follows because if P & R is true at all in a history it must be true at that time in the history when P is true. So the times at which condition P & R holds (assuming that there are some) will be a subset of the times when condition P holds. Whatever is conditionally obligatory given P (or FP) for s is true in each scenario of s's ought set at those times when P holds. But then it is also true in each member of s's ought set at those times when P & R holds, and hence is conditionally obligatory given P & R (or F(P & R)).

It is also troubling that none of (99) - (101) provides a suitable symbolization and analysis for a conditional obligation to do Q given not-P. In Chisholm's puzzle, for example, we wish to be able to express a man's obligation not to tell his neighbors he is coming if he does not go to help them. But I see no analysis analogous to (101) for such a sentence. The condition, that he does not go, is not one which, if ever true in a history, is true at most once in a history. There are instead many times in each history when it is true that he does not go. Indeed, if his going to help his neighbors describes an instantaneous individual event, then in each history it is true at every time in that history, except perhaps one, that he does not go. And it seems much too strong to require that the agent does not tell in each history in each ought set of <u>all</u> those times, when the conditional obligation sentence is true. Even if the English sentence "he goes to help his neighbors" is taken to describe an event which is true at most once in each history, it is difficult to see how "he does not go to help his neighbors" describes an event with the same feature.

One might believe that if P is a sentence describing an event true at most once in a history then drawing a distinction between not doing P and refraining from P will solve this difficulty. For one might argue that refraining from P is an event true at most once in a possible future, since refraining from P requires consciously considering performing P and yet purposively not doing P.

However there are at least two reasons why this strategy will not solve our problem, even if a clear distinction can be made between not doing P and refraining from P. First, our analysis for conditional obligation sentences with negated conditions will then be restricted to conditions which are negations of acts performed by agents rather than events in general. Second, if refraining from P requires conscious thought and effort on the part of the agent, then the analysis will not even apply to all conditions which are negations of acts of agents. For example, the analysis will not be applicable to failures to do P which are instances of thoughtless negligence.

Amended versions of (99) - (101) alone are insufficient for analyzing all the conditional obligation sentences we wish to express and yet I see no way of developing analogous analyses for sentences with complex conditions.

Furthermore, although we carefully formulated (101) and (102) to allow expressions of conditional obligation sentences in which what is conditionally obligatory is to occur prior to the condition, nevertheless it turns out that interesting conditional obligation sentences (in particular, those sentences describing cases where fulfillment of the conditional obligation is not inevitable) with this temporal relationship will not be true given (102) after all.

To see this, consider Chisholm's second sentence, interpreted as expressing a certain man's conditional obligation to tell his neighbors he is coming given that he goes to their assistance, where the telling is to occur before the going. According to (102) this conditional obligation is true at α just in case (*i*) there is a history in which after α he goes to help his neighbors and tells them he is coming and (*ii*) every time β, $\alpha < \beta$, in a history through α and β, at which he goes to help his neighbors is such that every history morally acceptable at β has in it a time γ, $\alpha < \gamma < \beta$, at which he tells them he is coming.

Although there are many alternative <u>futures</u> at a time β, some of which are morally acceptable, there is only one past with respect to β. If we rule out moral blind alleys, as Thomason does, and require that at every time β there is at least one morally acceptable history, then the unique past with respect to β is part of that history and so is morally acceptable at β.

Thus if

(98) $O_{s'}(FG/FT)$

is true at α, then according to our analysis, every time β, $\alpha < \beta$, in a history through α and β at which G is true is such that there is an earlier time γ in that history, $\alpha < \gamma < \beta$, at which T is true. That is, if the conditional obligation is true, then every history in which he goes is a history in which he tells as well. No history in which G is true at β_1, say, is such that there is no γ_1 on that history, $\alpha < \gamma_1 < \beta_1$, at which T is true. Yet in our example it is not impossible at α to go without telling; at α it <u>is</u> open to the agent to go and tell, to go and not tell, to not go and tell, and to not go and not tell. At α there surely <u>are</u> possible futures in which he goes but does not tell. Thus $O_{s'}(FG/FT)$ is not true at α given our analysis, although the corresponding English sentence is true.[28]

It is natural to suggest that these difficulties which arise for the proposals we have been discussing could be avoided if, instead of focusing on <u>every</u> time at which the condition is true, we merely look at <u>one</u> time at which the condition holds. And it is natural to believe that this latter suggestion could be developed if we appealed again to the counterfactual conditional expression, for formulas A and B, A > OB, which would force us to focus on a single time in a history at which condition A is true.

It seems to me that there are three major considerations which motivate the use of the counterfactual conditional for expressions of conditional obligation sentences, if these sentences are to be analyzed within Thomason's branching time model. First, if A > OB directs us to look at a particular time in a history at which A is true and then determine whether or not B is true in the agent's ought set at that time, then conditional obligation sentences in which the time of the obligation is earlier than the time of the condition may no longer be problematic. For even if the agent tells his neighbors he is coming in the history before that time which the counterfactual picks out when he goes to help them, there may be other future times at which he goes to help them but before which there is no time when he tells them he is coming. Truth of the conditional obligation sentence will not restrict the possible alternative futures.

Second, our main argument in Chapter II against using $p > Oq$ as a general symbolization for conditional obligation sentences was that $p > Oq$ implies $p \supset Oq$. Thus if $p > Oq$ expresses a true conditional obligation for q given p and p is true, we must accept the general factual detachment of Oq. This argument is correct. Yet if the subjunctive conditional is built into Thomason's model this argument will no longer apply. For even if it follows that

$$(113) \quad V_{\alpha}(A > OB) = T \ \& \ V_{\alpha}(A) = T \supset V_{\alpha}(OB) = T,$$

this will not produce unwelcome results. If A is a purely present or past tense formula, detachment of the unconditional obligation seems unproblematic. Also, if A is a true future tense formula, FC, then C is true in every future scenario. And when C is inevitable, detachment seems acceptable as well. Whereas if A is a future contingent statement, $V_{\alpha}(A)$ will be undefined rather than true and so no detachment will be possible.[29]

And third, the counterfactual conditional expression $A > OB$ does not imply $A \ \& \ C > OB$.[30] No other analysis of conditional obligation we have evaluated, except Lewis', has allowed conditional obligations to vary with augmented conditions. And we have seen that Lewis' theory allows certain conditional obligations to follow given conjunctive conditions, even in cases where it seems that the added condition actually overrides the original conditional obligation.

It turns out that the task of introducing subjunctive conditionals into the branching time model is a particularly difficult one. Thomason has proposed one method in an unpublished paper, "A Theory of Conditionals in the Context of Branching Time," where he argues that if conditionals are added to the theory then the basic notion of the semantic theory, $V_{\alpha}^{h}(A)$, should be replaced. For, he says, the theory must be able to take account not merely of a future course of events, h, that is tentatively designated the actual one for the moment in which we find ourselves, but of "counterfactual actual futures" as well.

Although I shall not pursue Thomason's thoughtful and detailed suggestions for revising the theory to include conditionals, I shall discuss a difficulty which must be resolved if A > OB is used within Thomason's theory to symbolize sentences of conditional obligation. Furthermore I shall point out several consequences shared by any theory of conditional obligation built on the indeterminist time model and Thomason's theory of obligation.

If A > OB is used to symbolize conditional obligation sentences, then presumably we shall require a selection function to pick out a particular time at which A is true and a future scenario through that time (to evaluate the truth of A if A contains future tense operators). Whatever is true in the ought set at the selected time is then crucial for determining the truth of the conditional obligation sentence. We could perhaps build selection functions into the model structure or, as Thomason does, treat them as valuations. But neither formal development explains what makes the time selected a proper choice for deontic evaluation. If we are not going to look at all times when the condition holds, but will choose only one, what criteria can possibly mark off that time as the relevant one for deontic purposes?

If the semantic analysis of counterfactual conditional statements within the branching time model is analogous to Stalnaker's original semantic theory for the subjunctive conditional, then it would be misleading to say that the selected time is the nearest or closest or earliest time at which the condition holds.[31] For there is no ordering of times in our model analogous to the ordering of possible worlds based on comparative similarity posited by Stalnaker. Perhaps the selected time is to be that time such that if A were true it would be true then. But what reason do we have for believing that this time will be the appropriate time to examine the ought set? Furthermore, if the condition is a future tense statement or a negation, then it may be true at many times in the selected history and it is difficult to see how we can select a <u>particular</u> time such that if $\sim P$ or FP, for example, were true, it would be true then. Or, even if we can pick a particular time with this feature it is difficult to see why we should focus on a time at which FP is true rather than one at which P is true, at least in deontic contexts, where the time selected is the one at which we examine the ought set to determine the truth of the conditional obligation sentence. These puzzling questions must be answered if $A > OB$, interpreted within Thomason's model, is to serve as an acceptable symbolization for conditional obligation statements.

Moreover, the use of truth-value gaps leads to a peculiar consequence for any theory of conditional obligation built on the indeterminist time model. In providing a symbolization for Chisholm's sentences, even if we can develop an adequate formalization for the second and third sentences, interpreted as statements of conditional obligation, we may puzzle over how to symbolize the fourth sentence.

Suppose we agree that given Thomason's analysis of O, we should symbolize the first sentence of Chisholm's puzzle as

(1b) $O_{s'}FG$,

where s' names the man and G corresponds to "s' goes to help his neighbors." It is consistent with the truth of (1b) at α to symbolize the fourth sentence as

(4b) $\sim G$,

where (4b) is also true at α. But the truth of $\sim G$ at α, $V_\alpha(\sim G) = T$, merely tells us that s' does not go at α, and so does not adequately capture the meaning of the English sentence. A more natural reading of Chisholm's fourth sentence tells us that the man <u>will</u> not go to help his neighbors.

But the alternative symbolization

(4c) $F\sim G$

does not express the English sentence either. For (4c) is true at α just in case

$$(114)\quad (h)(h \in H_{\alpha} \supset (\exists\beta)(\beta \in h \ \&\ \alpha < \beta \ \&$$

$$V^{h}_{\beta}(\sim G) = T)).$$

And (114), though true, is uninteresting. Even in those histories where there is a time later than α at which he goes, there is a time after α at which he does not go.

Thus we might symbolize the fourth sentence as

$$(4d)\quad \sim FG.$$

But if (4d) is true at α, then

$$(115)\quad (h)(h \in H_{\alpha} \supset (\beta)(\beta \in h \ \&\ \alpha < \beta \supset$$

$$V^{h}_{\beta}(G) = F)),$$

that is, there are no future scenarios open at α in which s' goes. And this contradicts the truth of (1b) at α, that s' goes in the morally acceptable futures open at α. Thus within the indeterminist time model we must interpret the fourth sentence as asserting that the agent will not go where this is a future contingency. In other words, we formalize the fourth sentence as

(4d) $\sim FG$

where the value of (4d) at α is undefined.

These symbolizations,

(1b) $O_{s'}FG$

and

(4d) $\sim FG$

when valuated at a time α as true and undefined, respectively, are not inconsistent. That s' goes to help his neighbors in all future histories that are morally acceptable at α is consistent with the truth that he goes in some future scenarios but does not go in others open at α. Furthermore, whatever analysis and symbolization we provide for the second and third sentences, we can avoid the contradiction spotted by Chisholm. Since the value of $\sim FG$ is undefined, we can not detach an obligation not to tell them he is coming.

Yet symbolizing the fourth sentence as a statement with undefined truth value is unsettling. First, it does not provide us with a formalization of the sentences for which all are true when valuated at the same time.

Second, and more worrisome, is the fact that it is natural to believe that the first and fourth sentences tell us that s' does go to help his neighbors in the morally acceptable futures open at α, and yet in fact he does not go. He does not go in the actual history; the actual future at α is not a member of his ought set at α. However, on Thomason's view, within the indeterminist time model we can not express this compelling interpretation of the sentences. Given his semantic theory we can not say that just one of the possible futures open at α is the one which will be realized without abandoning the point of the model. If a single possible future is the actual one with respect to α, then it is not clear what relation those times in the non-actual futures can bear to α.[32]

And third, although the lack of truth-value blocks factual detachment for a deliberative ought version of the puzzle, it is not clear how an analogous past tense or judgmental ought version of Chisholm's paradox could be avoided.

Two other major questions arise for any theory of conditional obligation built on Thomason's model and theory of obligation. First, Thomason developed his semantics for the deliverative use of "ought" so that a version of Kant's "ought implies can" principle does hold.

He says, "if we rule out moral blind alleys by requiring that for all α there is at least one h such that αSh, any formula having the form $OA \supset \sim L \sim A$ is valid."[33] If $MA = \sim L \sim A$, then M corresponds to an important use of "can," which Thomason suggests has these properties:

> (1) if we say a thing can happen, its subsequent happening shows that what we said was true;
>
> (2) the fact that in circumstances similar to ours a certain thing has happened is prima facie evidence for the claim that it can happen (supposing it to be in question whether it can happen).[34]

But if the alternative future histories we are considering are those which are temporally possible, or logically possible at a time α, then it may be argued that the version of the "ought implies can" principle which will hold for his concept of obligation, interpreted deliberatively, and thus any analogous deliberative concept of conditional obligation as well, is too broad. Let us return to Thomason's example (quoted on page 138) according to which it is true that I ought to give you $50 tomorrow and so every scenario in my ought set is such that on that scenario I do give you the money tomorrow. Then, as Thomason aptly describes,

> It is compatible with this situation that there are also scenarios ahead of me on which, through no fault of mine, this ought is cancelled. For instance, it may be a matter of chance whether I will be hit by a car later today and spend tomorrow in a coma in some hospital. At a time after the accident, on a scenario on which there _is_ such an accident, there will be no scenarios on which I give you $50 tomorrow, and so my ought set at the time won't contain any such scenarious. . . It is characteristic of such a theory (and might lead some to entertain alternative theories) that when I ought to pay you $50 tomorrow it will then also be true that I ought not to have an accident that will prevent me from paying you; and that I ought not to have a heart attack; and that I ought not to have a death in the family that will call me out of town suddenly. What makes _these_ implausible is that such matters are, to a large extent, out of my control, and when I promise to pay you $50 tomorrow it seems I am not promising (or even obligating myself) not to have a heart attack beforehand. The theory I'm proposing denies this; I want to say that oughts are _risky_, and if I ought to pay you the $50 tomorrow I ought not to have a heart attack that will prevent me from paying you. Nevertheless, I agree, it would be peculiar to say "I ought not to have a heart attack" in these circumstances--but that isn't because it's false.[35]

Even worse for Thomason's theory than the consequences he lists above, if there are scenarios ahead of me on which there is a flood and I subsequently do not pay you the $50, we will have to say that when it is true that I ought to see to it that there is no flood. This begins to seem ludicrous.

It appears that Thomason's theory of obligation and any theory of conditional obligation built upon it will make too much obligatory in the deliberative sense. The notion of possibility built in to the theory is too broad. What is obligatory in the sense of ought-to-do should be possible in the sense of being in the agent's control or power.[36]

An attempt to restrict an agent's ought set at a time in some way to limit what will be obligatory or conditionally obligatory seems impossible. However we have already recognized that English sentences such as "I pay you Saturday at 3:00" can not be symbolized in Thomason's language as it is given. Thus we have agreed to exclude sentences with references to clock-times from the domain of sentences which atomic statement variables may symbolize. Similarly, we might add an informal condition to further restrict the set of sentences which may be symbolized as atomics for true obligation and conditional obligation sentences to be of the form "x does P," where x names an agent and P names an action in x's control or power, for some suitable notion of "being in the agent's power." This limitation would be inappropriate for purely tense theoretic contexts, and is too restrictive if applied to conditions in conditional obligation sentences, but it would be appropriate for whatever is obligatory or conditionally obligatory.

Finally, a major issue which we have pointedly avoided discussing thus far is the question of how the relation S_s is to be explained. How do we determine which histories will be in an agent s's ought set at a time α? A formal semanticist may escape the issue by pointing out that it is a substantive ethical question. Thomason himself makes no proposals; he leaves the characterization of morally acceptable histories open. But to show that his theory of obligation and any conditional obligation proposal built on it are non-circular we must provide an independent specification of an agent's ought set at a time.

If David Lewis is correct that possible worlds can be ranked according to some standard of goodness and that for any obligation there is a threshold beyond which the better and better worlds are such that whatever is obligatory is true in them, then it would seem that similar reasoning would guarantee that possible alternative futures could be ranked according to some criterion of goodness and for any agent's obligation there would be a threshold beyond which the better and better histories would all be such that he fulfills his obligation in them. The better and better histories over the threshold would be in the agent's ought set at the time of comparison. In other words, then, we might view an agent's ought set at a time as the set of best histories open to him at that time, where the criterion of goodness is left unspecified, as it was for Lewis.

But Lewis' view that there is such a threshold is exactly what we have, agreeing with Goldman and McMichael, denied. If we are correct, then within Thomason's theory any characterization of an agent's ought set at a time as the set of best histories open at that time is doomed to fail. And an alternative characterization must be provided or other revisions imposed upon his theory to develop an analysis of conditional obligation built within his indeterminist time model.

In conclusion, careful examination of what appeared to be a plausible intuitive idea upon which to develop a theory of conditional obligation within Thomason's indeterminist time model indicated that depsite its apparent suitability for contrary-to-duty conditionals, it could not be formulated within the semantic theory. We discovered that an analysis which focuses on a single time at which the condition holds might fare better. In particular, there are several reasons to pursue the investigation of one such analysis employing the subjunctive conditional connective built within Thomason's branching time model.

However it is difficult to see how any analysis of that type can justify the crucial choice of any one time in a history at which the condition holds as suitable for deontic evaluation.

Furthermore, we have discussed several issues which must be resolved if Thomason's analysis of obligation and any analogous analysis of conditional obligation is to be satisfactory. We must pursue the suggestions made in Appendix II, since the original semantic theory provides no way of expressing obligations and conditional obligations which are to be fulfilled at a specific time. We must admit that within the theory Chisholm's four sentences may be symbolized so that none is false, although they can not be symbolized consistently as all true at the same time. Moreover, we must determine whether or not a past tense version of the paradox will arise. We must attempt to restrict the theory or adapt it so that obligations and conditional obligations are those acts within an agent's power, and should suggest reasonable characterizations of an agent's ought set at a time.

These issues are not trivial, and their multitude may lead many to seriously question the ultimate suitability of the indeterminist time model for deontic contexts. Nevertheless, our arguments have shown that it is unlikely that an adequate theroy of conditional obligation can be developed without a semantic theory which countenances the concept of an open future.

NOTES TO CHAPTER V

1. "Indeterminist time and truth-value gaps," Theoria 36 (1970), p. 265. When this comment is taken out of context it appears misleading for Thomason to suggest that the linear conception of time can countenance alternative futures merely as epistemic possibilities. For the concept of an open future may be viewed as alternative possible courses of events taking place in a single, unique course of time. In this case a linear conception of time provides a framework for a multiplicity of possible future courses of events which are not merely epistemically possible. Prior suggested, and Thomason is developing, a rival view according to which the concept of an open future is provided by reference to the nature of the course of events occurring in time. On this view we regard branching points as in the course of time itself. The former view, according to which events branch in time does not entail the latter, according to which there is a branching of time. (See Rescher and Urquhart, Temporal Logic, Springer-Verlag, New York 1971, especially pp. 70-74, and N. Rescher's "Truth and Necessity in Temporal Perspective" in R.M. Gale, ed., The Philosophy of Time, New York, 1967, for a discussion of these two approaches to the concept of an open future.) These two views lead to different semantic theories. And in the context of Thomason's project, to provide a semantics for tense logic for which every formula has a truth value at a time, the model of indeterminist time allows, while the model of linear time does not, development of the concept of an open future. The indeterminist time model may provide the most helpful model for a semantics for tense logic, as Thomason suggests. How appropriate it is for deontic contexts is an issue we hope to clarify in this chapter. (See below especially pp. 138-145 and Appendix II.)

2. Ibid., pp. 264-5.

3. Ibid., p. 281 and "Deontic Logic and Freedom of Choice in Moral Deliberation," p. 6.

4. Thomason, "Deontic Logic and Freedom of Choice in Moral Deliberation," pp. 5-6.

5. Hector-Neri Castañeda provides a useful explanation of this distinction in "On the Semantics of the Ought-to-Do" in Semantics of Natural Language, Davidson and Harman, eds., Dordrecht, 1972, especially pp. 675-678.

6. Thomason, "Deontic Logic as Founded on Tense Logic," p. 12.

7. Ibid., p. 9.

8. Ibid., p. 9.

9. These conditions are weaker than they may at first appear. Thomason's model is general and may include unconnected branching structures. But there should be no difficulty if we wish to further restrict the ordering, perhaps by requiring that the structure be connected, that is, that $(\alpha)(\exists\beta)(\beta < \alpha)$ & $(\alpha)(\beta)(\exists\gamma)(\gamma < \alpha \;\&\; \gamma < \beta)$ or there is a unique α, (the root) such that $\sim(\exists\gamma)(\gamma < \alpha \;\&\; (\beta)(\beta \neq \alpha \supset \alpha < \beta)$.

10. Thomason, "Deontic Logic as Founded on Tense Logic," p. 11.

11. Cf. ibid., pp. 11-12.

12. For the atomic case, $V_\alpha(OP \equiv P) = T$

$$\text{iff } (h)(h \in H_\alpha \supset V^h_\alpha(OP \equiv P) = T)$$

$$\text{iff } (h)(h \in H_\alpha \supset [V^h_\alpha(OP) = T \equiv V^h_\alpha(P) = T])$$

$$\text{iff } (h)(h \in H_\alpha \supset [(g)(\alpha Sg \supset V^g_\alpha(P) = T \equiv V^h_\alpha(P) = T]).$$

Whenever P is atomic the value of P at a time in a history is defined as the value of P at that time (see (78), p. 133) and so the above follows

$$\text{iff } (h)(h \in H_\alpha \supset [(g)(\alpha Sg \supset V_\alpha(P) = T) \equiv V_\alpha(P) = T]),$$

which is clearly true.

When A is truth functional yet has no future tense operators, the result follows similarly since $V^h_\alpha(A)$ and $V^g_\alpha(A)$ are determined with respect to the values of their atomic parts, and the latter are determined solely with respect to α.

However for formulas FA, $V_\alpha(OFA \equiv FA) = T$

iff $(h)(h \in H_\alpha \supset V^h_\alpha(OFA \equiv FA) = T)$

iff $(h)(h \in H_\alpha \supset [V^h_\alpha(OFA) = T \equiv V_\alpha(FA) = T])$

iff $(h)(h \in H_\alpha \supset [(g)(\alpha Sg \supset V^g_\alpha(FA) = T) \equiv V^h_\alpha(FA) = T])$

iff $(h)(h \in H_\alpha \supset [(g)(\alpha SG \supset (\exists\beta)(\alpha < \beta \; \& \; V^g_\beta(A) = T \equiv$
$(\exists\gamma)(\alpha < \gamma \; \& \; V^h_\alpha(A) = T])$.

Here there are two cases. The biconditional holds from right to left. If every history through α is such that A is true at a later time in that history, then every morally acceptable history through α has in it a later time at which A is true. But the biconditional may not hold from left to right. A may be true at a time after α in every morally acceptable history through α but not in every history through α. Thus $V_\alpha(A \supset OA) = T$ for any A although $V_\alpha(OA \supset A) = T$ only if A is atomic or has no future tense operators in it.

13. Cf. von Wright, "Deontic Logic and the Theory of Conditions," in Hilpinnen, ed., Deontic Logic: Introductory and Systematic Readings, pp. 159-160, for his emphasis of this point in a comparison of alethic modal and deontic logic.

14. See below, p. 115.

15. Thomason, "Deontic Logic as Founded on Tense Logic," pp. 1-6.

16. Ibid., p. 3.

17. Ibid., note 3, p. 4.

18. Thomason, "Deontic Logic and Freedom of Choice in Moral Deliberation," pp. 11ff.

19. Ibid., p. 11.

20. Cf. ibid., p. 11.

21. I do not believe that this proposal or its subsequent variants are proposals which Thomason himself would endorse, although they are based on his model. He would prefer, I believe, a proposal employing his obligation operator and the subjunctive conditional

connective. (See "Deontic Logic as Founded on Tense Logic," p. 1, cited in Chapter II, p. 37.) But given our arguments in Chapters II and III we shall begin by developing a dyadic conditional obligation operator. And it may be that a proposal of the type Thomason would prefer will be susceptible to many of the difficulties which we shall see arise for our present proposal.

22. This argument turns on our symbolization of Chisholm's sentences as $O_{s'}(FG/FT)$. Yet it might be argued that $O_{s'}(FG/PT)$, where P is a past tense operator, or some other symbolization employing both future and past tense operators is more appropriate. But see below, pp. 166-167 and especially note 28.

23. For example we might let P,Q,R. . . . symbolize sentences describing non-repeatable, instantaneous generic events and instantaneous individual events as they are analyzed by Richard Montague in "On the Nature of Certain Philosophical Entities." According to Montague, these types of events may be identified with (very particular) properties of moments of time, and thus may be such that they can not occur at more than one time. See especially op. cit. pp. 182-188.

24. See also pp. 136-138 and especially note 12 in Chapter V for a different explanation of the same point.

25. We can not neatly say that what follows is $V_\alpha(FO_sFQ) = T$ since, given (101), Q need not be true merely in the future in each morally acceptable history.

26. If we wish to use the inevitability operator, L, we may express the same results. That is, we may derive

(110) $V_\alpha(O_s(P/Q)) = T \;\&\; V_\alpha(LP) = T \supset V_\alpha(O_sQ) = T,$

and the more interesting

(111) $V_\alpha(O_s(P/FQ)) = T \;\&\; V_\alpha(LP) = T \supset V_\alpha(O_sFQ) = T.$

For in the latter case, for example, given $V_\alpha(LP) = T$, then $(h)(h \varepsilon H_\alpha \supset V^h_\alpha(LP) = T)$, and so $(h)(h \varepsilon H_\alpha \supset (g)(g \varepsilon H_\alpha \supset V^g_\alpha(P) = T))$. This reduces to $(h)(h \varepsilon H_\alpha \supset V^g_\alpha(P) = T$, or, equivalently, $V_\alpha(P) = T$. Applying modus ponens to the truth conditions for $V_\alpha(O_s(P/FQ))$ we may conclude that $(f)(\alpha Sf \supset (\exists\gamma)(\alpha < \gamma \& f \varepsilon H_\gamma \& V^f_\gamma(Q) = T))$, that is, $V_\alpha(O_sFQ) = T$.

The semantic rule for L, like the rule for O, directs us to evaluate the truth of LA at α solely by determining the value of A at α. Thus if A contains no future tense operators,

$$(112) \quad (\alpha)V_\alpha(LA \equiv A) = T).$$

Hence (110) and (111) say no more than (108) and (109).

If we wish to express that A is inevitable in the sense that A is true sooner or later in every history, we must use the symbolism LFA. LFA is true at α just in case $(h)\{h \varepsilon H_\alpha \supset (g)(g \varepsilon H_\alpha \supset (\exists\beta)(\beta \varepsilon g \& \alpha < \beta \& V^g_\beta(A) = T))\}$, equivalently, $(g)(g \varepsilon H_\alpha \supset (\exists\beta)(\beta \varepsilon g \& \alpha < \beta \& V^g_\beta(A) = T))$, which says that $V_\alpha(FA) = T$. And so if $V_\alpha(O_s(FP/FQ) = T$ and $V_\alpha(LFP) = T$, it will also follow that every history is such that there is a time at which Q is true (sooner or later) in every member of s's ought set.

27. Theories built on this model will allow fewer detachments than Greenspan would wish, however, since truth in every possible future is a stronger notion than Greenspan's unalterability for an agent.

28. I believe this argument also shows why (97) would fail as our analysis of conditional obligation sentences even if we used a symbolization for Chisholm's second sentence (and others like it) different from $O_{s'}(FG/FT)$.

29. Cf. pp. 158-161.

30. See Chapter II, p. 48. This follows assuming the relevant axioms of Stalnaker's theory for the counterfactual conditional still hold when that connective is built in to the indeterminist time model.

31. But even if the selected time were the closest time at which the condition A holds, what reason would we have for believing that time to be the appropriate one for examining the ought set if A is, for example, FP or ∿P? What _is_ the closest time at which FP is true? And why would that time be more appropriate for deontic evaluation than the later time at which P is true?

32. For Thomason's comments see Thomason, "Indeterminist time and truth-value gaps," pp. 270-271 and note 9, p. 270. More needs to be said on this intriguing issue.

33. Thomason, "Deontic Logic as Founded on Tense Logic," p. 11.

34. Thomason, "Deontic Logic and Freedom of Choice in Moral Deliberation," p. 4.

35. _Ibid_., pp. 3-4.

36. In an unpublished paper, "Doing the Best We Can," Fred Feldman has focused on this feature of the ought-to-do. He develops an intriguing theory of conditional ought-to-do which avoids many of the difficulties that plague other theories of conditional obligation. His theory is based on the theory of ought-to-do presented in his "World Utilitarianism" in Lehrer, ed., _Analysis and Metaphysics_, Dordrecht, 1975, pp. 255-271. On his view the histories we consider in a "branching time model" are what Feldman calls "life histories," maximal chains of sequences of events in a given agent's power. Thus his theory is carefully constructed to avoid the difficulty associated with Thomason's theory that we have just discussed.

But an agent's obligations and conditional obligations can not be determined solely with respect to his morally acceptable ("optimific" on Feldman's utilitarian view) life histories, for he does not live in isolation; other occurrences and agents in the world must be taken into account. Thus Feldman, in order to provide a utilitarian analysis, needs the concept of a "life history world." He says on page 265 of his "World Utilitarianism":

> I assume that for every person, s, and time, t, and life history, K, then available to him, there is exactly one possible world that is the one that would exist if s were to live out K.

It is these "life history worlds" which Feldman claims can be ranked according to some standard of goodness, and which are used to determine the agent s's obligations and conditional obligations in a fashion similar to Lewis'. His theory is subject to the peculiarity discussed on pp. 111-112 in our evaluation of Lewis' theory. Moreover, Feldman's crucial assumption, quoted above, is implausible and thus diminishes the plausibility of the entire theory.

BIBLIOGRAPHY

1. Anderson, Alan Ross, _The Formal Analysis of Normative Systems_, (Technical Report No. 2, Contract No. SAR/Nonr-609 (16), Office of Naval Research, Group Psychology Branch), New Haven, 1956. Reprinted in _The Logic of Decision and Action_, Rescher (ed.), University of Pittsburgh Press, (1967), 147-213.

2. Anderson, Alan Ross, "On the Logic of 'Commitment,'" _Philosophical Studies_ 10 (1959), 23-27.

3. Anderson, Alan Ross, "Reply to Mr. Rescher," _Philosophical Studies_ 13 (1962), 6-8.

4. Anderson, Alan Ross, "Some Nasty Problems in the Formal Logic of Ethics," _Noûs_ 1 (1967), 345-360.

5. Åqvist, Lennart, "Good Samaritans, Contrary-to-Duty Imperatives, and Epistemic Obligations," _Noûs_ 1 (1967), 361-379.

6. Åqvist, Lennart, "Improved Formulations of Act Utilitarianism," _Noûs_ 3 (1969), 299-323.

7. Åqvist, Lennart, "Interpretations of Deontic Logic," _Mind_ 73 (1964), 246-253.

8. Åqvist, Lennart, "'Next' and 'Ought'; Alternative Foundations for von Wright's Tense Logic, with an Aplication to Deontic Logic," _Logique and Analyse_ 9 (1966), 231-251.

9. Åqvist, Lennart, "A Note on Commitment," _Philosophical Studies_ 14 (1963), 22-25.

10. Bergström, Lars, "Hintikka on _prima facie_ obligations," _Theoria_ 40 (1974), 163-165.

11. Bergström, Lars, "Reply to Professor Hintikka," _Theoria_ 41 (1975), 39.

12. Castañeda, Hector-Neri, "On the Semantics of the Ought-to-Do," in Davidson and Harman (eds.), _Semantics of Natural Language_, Dordrecht (1972), 675-694.

13. Castañada, Hector-Neri, "Ought, Time, and the Deontic Paradoxes," Journal of Philosophy 74 (1977), 775-791.

14. Castañada, Hector-Neri, Thinking and Doing, Dordrecht, 1975.

15. Chellas, Brian, "Conditional Obligation," in Stenlund (ed.), Logical Theory and Semantic Analysis, Dordrecht (1974), 23-33.

16. Chisholm, Roderick, "Contrary-to-Duty Imperatives and Deontic Logic," Analysis 24 (1963), 33-36.

17. Chisholm, Roderick, "The Ethics of Requirement," American Philosophical Quarterly 1 (1964), 147-153.

18. Feldman, Fred, "On Doing the Best we Can," unpublished paper.

19. Feldman, Fred, "World Utilitarianism," in Keith Lehrer (ed.), Analysis and Metaphysics, Dordrecht (1975), 255-271.

20. Føllesdal, Dagfinn and Hilpinnen, Risto, "Deontic Logic: An Introduction," in Hilpinnen, R. (ed.), Deontic Logic: Introductory and Systematic Readings, Dordrecht (1971), 1-35.

21. Goldman, Holly, "Dated Rightness and Moral Imperfection," Philosophical Review 85 (1976), 449-487.

22. Goldman, Holly, "David Lewis' Semantics for Deontic Logic," Mind 86 (1977), 242-248.

23. Goldman, Holly, "Doing the Best One Can," forthcoming.

24. Greenspan, P.S., "Conditional Oughts and Hypothetical Imperatives," Journal of Philosophy 72 (1975), 259-276.

25. Greenspan, P.S., "Derived Obligation: Some Paradoxes Escaped," Harvard PhD dissertation, 1972.

26. Greenspan, P.S., "Oughts and Determinism: A Response to Goldman," Philosophical Review 87 (1978), 77-83.

27. Greenspan, P.S., "Practical Reasoning and Deontic Logic: Some Footnotes in Reply to Castañeda," forthcoming in Journal of Philosophy.

28. Greenspan, P.S., "Wiggins on Historical Inevitability and Incompatibilism," Philosophical Studies 29 (1976), 235-247.

29. Hansson, Bengt, "An Analysis of Some Deontic Logics," Noûs 3 (1969), 373-398. Reprinted in Hilpinnen (ed.), Deontic Logic: Introductory and Systematic Readings, Dordrecht (1971), 121-147.

30. Hilpinnen, Risto (ed.), Deontic Logic: Introductory and Systematic Readings, Dordrecht, 1971.

31. Hintikka, Jaakko, "Comment on Professor Bergström," Theoria 41 (1975), 35-38.

32. Hintikka, Jaakko, "Deontic Logic and its Philosophical Morals," in Models for Modalities, Selected Essays by J. Hintikka, Dordrecht (1970), 184-214.

33. Hintikka, Jaakko, "On the Logic of Perception," in N. Care and R. Grimm (eds.), Perception and Personal Identity, Case Western Reserve Press (1969).

34. Hintikka, Jaakko, "Some Main Problems of Deontic Logic," in Hilpinnen (ed.), Deontic Logic: Introductory and Systematic Readings, Dordrecht (1971), 59-104.

35. Kordig, Carl R., "Relativized Deontic Modalities" in Anderson, Marcus, Martin (eds.), The Logical Enterprise, Yale University Press (1975), 221-233.

36. Ladd, John, "Remarks on the Conflict of Obligation," Journal of Philosophy 55 (1958), 811-819.

37. Lenk, Hans, "Varieties of Commitment," Theory and Decision 9 (1978), 17-37.

38. Lewis, David, Counterfactuals, Cambridge: Harvard University Press, 1973.

39. Lewis, David, "Reply to McMichael," Analysis 38 (1978), 85-86.

40. Lewis, David, "Semantic Analyses for Dyadic Deontic Logic," in Stenlund (ed.), Logical Theory and

Semantic Analysis, Dordrecht (1974), 1-14.

41. McMichael, Alan, "Too much of a good thing: a problem in deontic logic," Analysis 38 (1978), 83-84.

42. Montague, Richard, "Logical Necessity, Physical Necessity, Ethics, and Quantifiers," Inquiry 4 (1960), 259-269.

43. Montague, Richard, "On the Nature of Certain Philosophical Entities," The Monist 53 (1969), 159-194. Reprinted in Formal Philosophy: Selected Papers of Richard Montague, Yale University Press (1974), 148-187.

44. Powers, Lawrence, "Some Deontic Logicians," Noûs 1 (1967), 381-400.

45. Prawitz, Dag, "The Alternatives to an Action," Theoria 36 (1970), 394-449.

46. Prior, A.N., "The Paradoxes of Derived Obligation," Mind 63 (1954), 64-65.

47. Purtill, R.L., "Deontically Perfect Worlds and Prima Facie Obligations," Philosophia 3 (1973), 429-438.

48. Rescher, Nicholas, "An Axiom System for Deontic Logic," Philosophical Studies 9 (1958), 24-30.

49. Rescher, Nicholas, "Conditional Permission in Deontic Logic," Philosophical Studies 13 (1962), 1-6.

50. Rescher, Nicholas, "Semantic Foundations for Conditional Permission," Philosophical Studies 18 (1967).

51. Rescher, Nicholas, "Truth and Necessity in Temporal Perspective," in R.M. Gale (ed.), The Philosophy of Time, New York (1967).

52. Rescher, N. and Urquhart, Temporal Logic, Springer-Verlag, New York, 1971.

53. Robison, John, "Further Difficulties for Conditional Permission in Deontic Logic," Philosophical Studies 18 (1967), 27-30.

54. Robison, John, "Who, What, Where, and When: A Note on Deontic Logic," Philosophical Studies 15 (1964), 89-92.

55. Ross, W.D., Foundations of Ethics, Oxford, 1939.

56. Ross, W.D., The Right and the Good, Oxford, 1930.

57. Sellars, Wilfred, "Reflections on Contrary-to-Duty Imperatives," Noûs 1 (1967), 303-344.

58. Stalnaker, Robert C., "A Theory of Conditionals," in Rescher (ed.), Studies in Logical Theory (American Philosophical Quarterly Supplementary Monograph Series), Oxford (1968), 98-112.

59. Thomason, Richmond H., "Deontic Logic and Freedom of Choice in Moral Deliberation," unpublished manuscript. Read at the American Philosophical Association Pacific Division Meetings, 1977.

60. Thomason, Richmond H., "Deontic Logic as Founded on Tense Logic," unpublished manuscript, (1974).

61. Thomason, Richmond H., "Indeterminist time and truth-value gaps," Theoria 36 (1970), 264-281.

62. Thomason, Richmond H., "A Theory of Conditionals in the Context of Branching Time," unpublished manuscript.

63. van Fraassen, Bas, "The Logic of Conditional Obligation," Journal of Philosophical Logic 1(1972), 417-438.

64. van Fraassen, Bas, "Values and the Heart's Command," Journal of Philosophy 70 (1973), 5-19.

65. von Wright, Georg Henrick, "A Correction to a New System of Deontic Logic," Danish Yearbook of Philosophy 2 (1965), 103-107. Reprinted (in part) in Hilpinnen (ed.), Deontic Logic: Introductory and Systematic Readings, Dordrecht (1971), 115-120.

66. von Wright, Georg Henrick, "Deontic Logic," Mind 60 (1951), 1-15. Reprinted in Logical Studies by G.H. von Wright, London: Routledge and Kegan Paul (1957), 58-74. Also reprinted in Contemporary Studies in Logical Theory, Copi and Gould (eds.), New York: MacMillan (1970), 303-315.

67. von Wright, Georg Henrick, "Deontic Logic and the Theory of Conditions," in Hilpinnen (ed.), Deontic Logic: Introductory and Systematic Readings, Dordrecht (1971), 159-177.

68. von Wright, Georg Henrick, "Deontic Logics," American Philosophical Quarterly 4 (1967), 136-143.

69. von Wright, Georg Henrick, "A New System of Deontic Logic," Danish Yearbook of Philosophy 1 (1964), 173-182. Reprinted in Hilpinnen (ed.), Deontic Logic: Introductory and Systematic Readings, Dordrecht (1971), 105-120.

70. von Wright, Georg Henrick, "A Note on Deontic Logic and Derived Obligation," Mind 65 (1956), 507-509. Reprinted in Contemporary Studies in Logical Theory, Copi and Gould (eds.), New York: MacMillan (1970), 316-318.

APPENDIX I

Most systems of deontic logic include some form of a basic deontic logic as a subsystem. Bengt Hansson, in "An Analysis of Some Deontic Logics" (in Risto Hilpinnen, ed., Deontic Logic: Introductory and Systematic Readings), has isolated this basic system and called it the standard deontic logic (SDL). Since then SDL has been referred to elsewhere in the literature and serves as a standard system within which it is possible to express most of the philosophical problems associated with deontic logic.

According to Hanson SDL is based on a basis logic, BL, which may be the propositional calculus or any related system, provided valuations and validity are defined in the usual way and provided BL is complete, that is, every valid formula is a theorem. What constitutes a well formed formula and a theorem is determined in the usual way.

SDL is minimal in the sense that it is the smallest set fulfilling the requirements that

(i) whenever f is a formula of BL then Of is a formula of SDL,

(ii) the negation of any formula of SDL is a formula of SDL,

and (iii) the disjunction of any formulas of SDL is a formula of SDL.

The deontic operator O is read "it is obligatory that". F, read "it is forbidden that", and P, read "it is permissible that', may be defined such that

(iv) $Fp \equiv O(\sim p)$

and (v) $Pp \equiv \sim Fp \equiv \sim O(\sim p)$

where p is a schematic sentence letter expressing a possible act or state of affairs. Truth functional formulas of SDL are defined as usual but no iterated operators (e.g., OOp) nor mixed formulas (e.g., $p \vee Op$) are allowed in the language.

Every formula which is obtained from a theorem of the propositional calculus by substituting formulas of SDL for the variables is an axiom of SDL. More than one set of deontic axioms might serve as the deontic base for SDL. One set suggested by Hansson and used by others is:

(1) $Op \supset \sim O(\sim p)$

and (2) $O(p \supset q) \mathbin{\&} Op \supset Oq$ (equivalently,

$O(p \supset q) \supset (Op \supset Oq)$).

Furthermore, many authors have found it reasonable to add

(3) O(t),

where t is a tautology in BL, since denial of (3) excludes only empty normative systems or those cases in which nothing is obligatory (cf. Føllesdal and Hilpinnen, p. 13). Hansson points out that (3) is sometimes added not as an axiom but as an inference rule,

(R) If a formula f is provable in BL,
then Of is provable in SDL,

along with modus ponens and a suitable substitution rule.

This system of SDL is equivalent to the original system proposed by von Wright in "Deontic Logic" in Mind, 1951, in which permission is the deontic primitive and the axioms are thus formulated in terms of the deontic operator P, except that von Wright prefixed deontic operators to names of acts and did not propose (3). Hansson suggests that although von Wright did not propose (3), nevertheless the system SDL as presented above, including (3) or (R), is essentially what von Wright intended. Thus von Wright truly did provide the foundation for deontic logic.

APPENDIX II

We wish to investigate the possibility of preserving Thomason's system and (1) providing a symbolization of sentences in the form of (89), (2) maintaining the idea underlying his informally stated principle (see (94)), and (3) avoiding the difficulty presented in Chapter V on pages 138-145. Suppose we enrich (no pun intended) his language by adding to it a way of referring to "clock-times" such as Saturday at 3:00, and argue that sentences such as "I pay you Saturday at 3:00" can not be symbolized as atomic sentences in the language. We shall then say that although "I pay you" may be formalized as an atomic sentence, any reference to a clock-time such as Saturday at 3:00 must be made explicit and can not be embedded within a symbolization for an atomic sentence. We have discovered that clock-times such as Saturday at 3:00 can not be symbolized in Thomason's language as it is presented and thus we must augment his language to enable explicit reference to clock-times across histories.

Interestingly, in adapting his semantic analysis for the deliberative use of "ought" to the judgmental use, Thomason argues that a similar cross reference may be appropriate. For in determining what ought to have been done (although it is now too late to do it) one imagines that at some point in the past things went more as they should than they in fact did. Thus one in effect transports oneself to a (not necessarily unique) instant, a set of counterfactual moments, if you wish, which ought to have been realized. Thomason says,

> ...suppose that I've drawn up to a stop line of cars at a red light and, looking in my rear mirror, I see a car coming up behind me too fast for there to be a chance of its stopping in time. There is a sense in which the driver of the car ought not to hit me. While I'm waiting there with nothing better to do, I make this judgment in the following way. First, I go back a few seconds to a point in time at which the driver still had a chance of stopping safely. This, as I've said, is an act of imagination: wishful imagination, perhaps. Then I consider a variety of alternative scenarios in which he drives as he ought. In general there is more than one of these... All of these scenarios are, of course, might-have-beens, since he didn't in fact drive as he ought. Along each of these scenarios, then, I choose a particular instant to serve as an alternative for the one in which I unhappily find myself. The most natural way of doing this in our example is to use the metric properties of time and take instants along the other scenarios in which clocks show the same time they do at the instant in which I find myself. ("Deontic Logic as Founded on Tense Logic", p. 13)

Our interest is not in developing a semantics for the judgmental use of "ought". However if we are to make explicit reference to clock-times, we need a way of choosing an instant along each future history or scenario corresponding to, for example, Saturday at 3:00. In the last sentence of the above quotation Thomason suggests a natural way of doing this.

Let us again suppose as we did in Chapter I that time, conceived of linearly, can be partitioned into discrete instants which we may refer to by definite descriptions, definite descriptions of what we have been calling clock-times. Then suppose we introduce to our language t_1, t_2, t_3,... as non-rigid designators for descriptions such as "Saturday at 3:00", such that if t_3 designates Saturday at 3:00, then for every history or scenario h there will be a β such that the value of t_3 at h is β, which we may abbreviate with a function $V^*(t_3,h) = \beta$. In other words the non-rigid designator t_3 picks out in each history an instant corresponding to Saturday at 3:00, perhaps the instant at which clocks and calendars in the history indicate it is Saturday at 3:00. (But see below, Appendix II, pp. 204-205.)

The idea of using non-rigid designators in this way to refer across histories or worlds is not new. Jaakko Hintikka made a similar suggestion in "On the Logic of Perception" (in Perception and Personal Identity, N. Care & R. Grimm, eds., Case Western Reserve Press, 1969), to refer to intensional objects in formalizations of certain perceptual statements.

Although our project need not include a complete explanation of the metric properties of clock-times, it does seem that we will preserve the metric structure of time if we require that for every t_i and t_j, $t_i < t_j$ if and only if the clock-time named by t_i is earlier than the clock-time named by t_j. Furthermore, it seems reasonable to require that for every t_i and history h there is one and only one β such that $V^*(t_i,h) = \beta$. On any history there is a β designating Saturday at 3:00, for example, and that β is unique.

Given this minimal explanation, we might now let atomic sentence variables P, Q, R,... stand for sentences like "I pay you", where no reference to a clock-time appears in the sentence, and P_{t_3} for "I pay you Saturday at 3:00". Then we wish to say that P_{t_3} is true at any time in a history just in case P is true at the time corresponding to Saturday at 3:00 in that history. More generally, we might formulate Thomason's principle for truth of a formula at a clock-time as:

$$(1) \quad (\alpha)(h)(V^h_\alpha\ (A_{t_i}) = T \text{ iff } V^h_{V^*(t_i,h)}(A) = T).$$

Because $V^*(t_i,h)$ is clearly a time in history h, we may simplify (1) to

$$(2) \quad (\alpha)(h)(V^h_\alpha\ (A_{t_i}) = T \text{ iff } V_{V^*(t_i,h)}(A) = T).$$

And finally, returning to the semantic rule for deliberative oughts we may have

$$(3) \quad V^h_\alpha\ (OA_{t_i}) = T \text{ iff } (g)(\alpha SG \supset V^g_\alpha\ (A_{t_i}) = T.$$

And given (1) and (2) this is equivalent to

$$(4) \quad V^h_\alpha\ (OA_{t_i}) = T \text{ iff } (g)(\alpha SG \supset V_{V^*(t_i,g)}(A) = T).$$

We now have a rich enough meta-language to give a semantic rule for expressions of obligations in which what is obligatory is to be brought about at a certain time. That is, we can distinguish the time at which the obligation is said to hold and the time at which it is to be fulfilled. (Cf. Castañeda, "Ought, Time, and the Deontic Paradoxes", p. 782, on the importance of this distinction.) When the obligation statement is interpreted deliberatively we do not wish, for example, to be obligated Tuesday to do A Monday, and we now have a language within which we can require that

$$(5)\quad V^{h}_{\alpha}\,(OA_{t_i}) = T \supset V^{*}(t_i,h) \geqslant \alpha.$$

(Compare (2), p. 21, of Chapter I.) And we may wish to incorporate this condition in (4) so that

$$(6)\quad V^{h}_{\alpha}\,(OA_{t_i}) = T \text{ iff } (g)(\alpha Sg \supset V^{*}(t_i,g) \geqslant \alpha \;\&\; V_{V^{*}(t_i,g)}(A) = T),$$

where, as noted in Chapter V, h is irrelevant. On this analysis, a sentence indicating an obligation, for example, to pay you Tuesday at 4:00 will be symbolized without the use of a future tense operator, but with a sentence variable and a non-rigid designator for Tuesday at 4:00.

There seem to me to be three main problems with this suggestion. First, although any formula can presumably be true at a particular moment or clock-time, t_i, it is very difficult to see how the time, t_i, at which a complex formula (a negation or material implication, for example) is true is related to the time or times at which its atomic parts are true. Thus it might be best to view (1) - (6) as principles which hold only when A is an atomic sentence letter. But then we have a much more restricted language than Thomason's and have prevented ourselves from providing a symbolization for the sentence expressing Jones' obligation not to tell his neighbors he is coming, for example.

Second, in stipulating that a non-rigid designator, t_i, picks out the instant in each history at which the clocks and calendars in that history register the same time and date, we are restricting the set of alternative possible futures. For we can not account for possible futures in which the clocks and calendars go haywire.

Third, if two times in different branches of the temporal model structure are truly incomparable in the sense that they are neither identical with each other nor is one earlier than the other, then it would seem that their incomparability precludes the very possibility of choosing non-rigid designators to pick out corresponding instants in each alternative history. If we may truly slice across histories to find clock-times then it appears we really are conflating our model, however covertly, to a conception of a unique course of time in which there are alternative possible courses of events. (See Chapter V, note 1.)

To the second objection we might reply that alternative futures in which the clocks go haywire are indeed irrelevant for the semantic evaluation of sentences expressing obligations which are to be discharged at a particular time. And in response to the third objection it must be pointed out that to pick out times from each history to correspond to a definite description is not to claim that they are identical, which would clearly violate their incomparability. Nor is the choice of times picked out by a particular non-rigid designator absolute in any sense. The semantic proposals given here merely attempt to provide a semantic analysis for a theory countenancing a single course of time within Thomason's more general semantic model.