INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)” . If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeab Road
Ann Arbor, Michigan 48106
EAKER, Charles Edward, 1940-
A DEFENSE OF MONADIC DEONTIC LOGIC.

The Ohio State University, Ph.D., 1973
Philosophy

University Microfilms, A XEROX Company, Ann Arbor, Michigan
A DEFENSE OF MONADIC

DEONTIC LOGIC

Dissertation
Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School in the Ohio State University

By

Charles Edward Eaker, B.A.

The Ohio State University
1973

Reading Committee:

Charles F. Kielkopf
William G. Lycan
Bernard Rosen

Approved by

Charles F. Kielkopf
Adviser
Department of Philosophy
ACKNOWLEDGMENTS

The major contributor to the form of my present views about deontic logic is my advisor, Charles Kielkopf, not because my views are his views—we disagree on some fundamental issues—but because of his knowledge of deontic logic and because of his patience and persistence in criticizing my arguments and helping me to see the force of the arguments of others. His encouragement has been unfailing, and he has always done everything in his power to handle administrative details, particularly at times when I had given up hope. My sincere thanks to him for all he has done.

The shape of the present work has also been influenced by discussions with two of my colleagues in the philosophy department at Oswego—Bob Carnes whose reluctance to abandon his positions has made my own views stronger than they otherwise would have been, and Ken Friedman who has an incredible ability to produce a counter-example at a moment's notice. Their ability and willingness to discuss virtually any topic at almost any time has been an invaluable asset.

But all of this would never have been without the constant encouragement, support, patience, understanding, and just plain nagging from Susan Eaker who, quite justifiably, has recently come to call this "our" dissertation. And I happily acknowledge the complete lack of cooperation from Erin Eaker during the entire two years of her life.
June 12, 1940 .......... Born - Detroit, Michigan

1964 ............... B.A., Capital University, Bexley, Ohio

1965-1969 .......... Teaching Assistant, Department of Philosophy, The Ohio State University, Columbus, Ohio

1969-1970 .......... Lecturer, Department of Philosophy, The Ohio State University

1970- .............. Assistant Professor, Department of Philosophy, State University of New York College of Arts and Sciences at Oswego
TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................ ii
VITA ....................................................... iii

Chapter

I. WHAT IS DEONTIC LOGIC?

1.1 In the Beginning ...................................... 1
1.2 Jørgensen's Dilemma .................................. 3
1.3 A Synopsis .............................................. 6

II. ALETHIC MODAL LOGIC AND POSSIBLE WORLD SEMANTICS

2.1 An Introduction ........................................ 10
2.2 Standard Alethic Logics .............................. 10
2.3 Truth Trees ............................................. 13
2.4 Interpretations of Alethic Logics ................. 15
2.5 Deontic Interpretations ............................. 21
2.6 When Is a Logic Deontic? ......................... 29

III. VON WRIGHT'S ORIGINAL DEONTIC LOGIC

3.1 Introduction ........................................... 38
3.2 The Formal Structure of P .......................... 39
3.3 The Interpretation of P .............................. 42
3.4 The Decision Procedure for P ..................... 47
3.5 The Principle of Permission ....................... 55
3.6 Some Theorems and Some Limitations .......... 61

IV. ANDERSON'S REDUCTION

4.1 Anderson's Proposal .................................. 65
4.2 Anderson's Reduction ............................... 68
4.3 The Systems OM, OM', and OM'' ................ 78
4.4 Some Theorems and a Defense ..................... 79
4.5 The Use of 'B' Criticized ......................... 81
4.6 A Look at the Sanction ............................. 82
## V. IDEAL WORLD SEMANTICS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Deontic Analogs of T, S4, and S5</td>
<td>90</td>
</tr>
<tr>
<td>5.2</td>
<td>Iterated Deontic Operators</td>
<td>92</td>
</tr>
<tr>
<td>5.3</td>
<td>Iterated Moral Oughts</td>
<td>95</td>
</tr>
<tr>
<td>5.4</td>
<td>The Paradoxes of the Standard Deontic Logics</td>
<td>101</td>
</tr>
<tr>
<td>5.5</td>
<td>Ross's Paradox</td>
<td>102</td>
</tr>
<tr>
<td>5.6</td>
<td>A Variation on Ross's Paradox</td>
<td>104</td>
</tr>
<tr>
<td>5.7</td>
<td>The Good Samaritan Paradox</td>
<td>106</td>
</tr>
<tr>
<td>5.8</td>
<td>The Derived Obligation Paradox</td>
<td>108</td>
</tr>
</tbody>
</table>

## VI. HINTIKKA, QUANTIFIERS, AND PRIMA FACIE OBLIGATIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Quantifiers</td>
<td>113</td>
</tr>
<tr>
<td>6.2</td>
<td>The Interpretation of HDL</td>
<td>119</td>
</tr>
<tr>
<td>6.3</td>
<td>Deontic and Logical Consequences</td>
<td>137</td>
</tr>
<tr>
<td>6.4</td>
<td>A Critique of HDL</td>
<td>143</td>
</tr>
<tr>
<td>6.5</td>
<td>Does Ought Imply Can?</td>
<td>153</td>
</tr>
<tr>
<td>6.6</td>
<td>The Contrary-to-Duty Paradox</td>
<td>158</td>
</tr>
</tbody>
</table>

## VII. CASTANEDA AND THE LOGIC OF AGENCY

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Imperatives and Obligation</td>
<td>162</td>
</tr>
<tr>
<td>7.2</td>
<td>The Formal Structure of CDL</td>
<td>169</td>
</tr>
<tr>
<td>7.3</td>
<td>The Semantics of CDL</td>
<td>172</td>
</tr>
<tr>
<td>7.4</td>
<td>A Critique of CDL</td>
<td>176</td>
</tr>
</tbody>
</table>

## VIII. PRELIMINARIES TO AN IMPROVED DEONTIC LOGIC

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>The Uses of 'Ought'</td>
<td>188</td>
</tr>
<tr>
<td>8.2</td>
<td>The Logical Requirements of Agency</td>
<td>195</td>
</tr>
<tr>
<td>8.3</td>
<td>The Logic of Conscious Action</td>
<td>198</td>
</tr>
<tr>
<td>8.4</td>
<td>Tree Rules for CL</td>
<td>202</td>
</tr>
<tr>
<td>8.5</td>
<td>An Imperative Logic</td>
<td>205</td>
</tr>
<tr>
<td>8.6</td>
<td>Quantifiers</td>
<td>211</td>
</tr>
</tbody>
</table>

## IX. AN IMPROVED DEONTIC LOGIC

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1</td>
<td>The Syntax of IDL</td>
<td>218</td>
</tr>
<tr>
<td>9.2</td>
<td>The Uses of 'Ought' Formalized</td>
<td>222</td>
</tr>
<tr>
<td>9.3</td>
<td>Dealing With Moral Exceptions</td>
<td>232</td>
</tr>
</tbody>
</table>

## BIBLIOGRAPHY

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>246</td>
</tr>
</tbody>
</table>
1.1 In The Beginning

For all practical purposes deontic logic was born in 1951 with the publication of two papers by G. H. von Wright in which he presented what is now known as a monadic deontic logic. Monadic deontic logics (as opposed to dyadic deontic logics) typically deal with expressions such as

(1) $O p$
(2) $P p$
(3) $F p$

which are read as 'p is obligatory', 'p is permitted', and 'p is forbidden', respectively. A dyadic deontic logic applies these same operators to more complex expressions:

(4) $O(p/q)$
(5) $P(p/q)$
(6) $F(p/q)$

which are read as 'p is obligatory given q', 'p is permitted given q', and 'p is forbidden given q'.

---

1von Wright [1951a] and [1951b]. (The author *cum* date references are to the bibliography which follows the text.)
After an extremely difficult childhood of about ten years, monadic deontic logics were abandoned in favor of dyadic deontic logics. This was precipitated by the discovery of paradoxical theorems found in all extant monadic deontic logics, the most worrisome being Ross's paradox,

\[(7) \quad Op \rightarrow O(p \lor q)\]

which can be read as 'If I ought to give this person aid, then I ought to give him aid or kill him.' Given this reading, (7) does not seem to be a desirable candidate for a deontically valid expression. Another theorem which is difficult to accept is known as the paradox of derived obligation:

\[(8) \quad Fp \rightarrow O(p \rightarrow q)\]

This one can be read as 'If I do something that is forbidden, then I should assassinate a public figure.' In Chapter V these paradoxes will be interpreted as viciously as possible. Even so, dyadic deontic logics which were designed to avoid these paradoxes are complicated and I find something aesthetically pleasing about the comparative simplicity of monadic deontic logics. Personal entropy more than anything else has kept me clinging to monadic deontic logics in the face of extremely strong pressures. Now, with a certain degree of sadness, I find it my duty to report that this inertia has been rewarded. My major thesis is that monadic deontic logics were not only abandoned in haste, but unnecessarily. I will argue that there is an interpretation of monadic deontic logic hitherto unnoticed by deontic logicians; an interpretation that successfully avoids the paradoxes. I will retain (7) and (8) as theorems, but I will insist
that they should be interpreted in a manner other than the one suggested above. Furthermore, I will attempt to show that the resulting monadic deontic logic provides a foundation for a deontic logic which goes even further.

The other side of this thesis is that there are several clearly distinguishable ways in which the word 'ought' is used in ordinary discourse. Some of these uses, incidentally, are not normative uses. I will identify several of these uses and show that they have different logics. Accordingly, dyadic deontic logics will not be discussed in the present work and the term 'deontic logic' will be treated as a shortened form of 'monadic deontic logic'.

1.2 Jørgensen's Dilemma

It is apparently significant that this dilemma is named after a Dane\(^2\). But I am ahead of myself. What is the dilemma? One horn is the claim that deontic sentences (indicative sentences in which 'ought' occurs) are neither true nor false. Since validity is defined in terms of these two categories, there are no logically valid arguments which contain deontic sentences, so forget about the possibility of deontic logic. The other horn notes that there are 'valid' (in some sense of the term) inferences which contain deontic sentences. For example:\(^3\)

\(^2\)This dilemma was named by Alf Ross. See Ross [1941].

\(^3\)The examples are taken from Jørgensen [1938], p. 290.
(9) Keep your promises!
This is a promise of yours.
Hence: Keep this promise!

(10) Love your neighbor as yourself!
Love yourself!
Hence: Love your neighbor!

Ah, hal you say, (9) and (10) contain only imperatives; there is not a deontic sentence to be seen. Quite true. Yet a native speaker of English finds it disconcerting to read Scandinavian treatises on deontic logic which uncritically treat directives, commands, imperatives and deontic logic as all being about the same thing. And here lies the appropriateness of the name. In a footnote, Alf Ross suggests that Danes probably do not see the difference between imperatives and oughts which we see because of an important linguistic difference between Danish and English. But that does not matter. Imperatives are certainly classic examples of sentences which have no truth value and any number of English speaking philosophers have claimed that deontic sentences are imperatives or, like imperatives, have no truth value. So, the dilemma stands, although I believe that it is stronger when applied to imperatives.

---

4 For example, Ross [1968] and Espersen [1967].
5 Ross [1968], p. 154.
6 Ayer [1936], p. 103.
7 Barnes [1934], p. 45.
There are several avenues of escape. A favorite of Scandanavians is to transform an imperative which has no truth value ("Close the door!") into an improper imperative ("It is commanded that the door is closed.") for logical purposes. One might develop an imperative semantics which uses such notions as "in force" and "not in force" or "satisfied" and "not satisfied" to define Imperative validity. Grasping the horns, one might deny the possibility of this kind of inference or, in the case of deontic logic, simply claim that deontic sentences have a truth value.

It is clear enough, I think, that

\[ (11) \quad \text{You ought to keep the promise or apologize.} \]
\[ \text{You ought not keep the promise.} \]
\[ \text{Hence: You ought to apologize.} \]

is a "good" deontic argument, and that

\[ (12) \quad \text{People kill for wealth.} \]
\[ \text{Hence: People ought to kill for wealth.} \]

is a "bad" deontic argument. Hence I cannot sympathize with any view which holds that deontic arguments are inherently impossible. And if it is clear that truth values may not be applied to the components of (11) and (12) (which it is not), then some other set of semantic values must be assigned to them and deontic "validity" defined in terms of

---

8 Jørgensen [1938], pp. 292-293. In addition, Ross and Espersen report that this view is taken up by Hedenius [1941] and Moritz [1954].

9 Bohnert [1946], Gombay [1964], Rescher [1966], and Sosa [1966].

10 Williams [1963] and Keene [1966].
those new values. Modern formal semantics is quite indifferent about how the semantic values are interpreted. (More on that in the next chapter.) So, imperative logic and deontic logic can be "done" at least in a formal way. Whether the results will strike an intuitive chord or not remains to be seen. I have little patience with a priori attempts to settle this question since it seems to be an empirical one. However, I know of no particularly strong reasons for supposing that deontic sentences do not have a truth value and since my natural inclination to suppose that they do is buttressed by the fact that normative discourse treats them as though they do, I shall freely apply truth values to deontic sentences—it is so much more convenient to do so. Besides, I don't know what other words to use.

1.3 A Synopsis

I shall close these introductory remarks with a brief synopsis of the sequel. First, I take up some necessary preliminaries. So much of the work in deontic logic borrows from alethic logic that a brief introduction to alethic logic seems indispensible. This introduction will do two things. It familiarizes the reader, even if ever so slightly, with three classic alethic modal logics, and it also allows me to introduce semantic techniques that have proved quite fruitful in the general area of modal logic and which will be used throughout the present work. It also permits the formulation of a set of criteria, the Criteria of Adequacy, against which the deontic logics we will look at may be judged. The criteria are not merely formal criteria.
I believe that a philosophically acceptable deontic logic must be able to help us clarify at least some of our normative concepts. To do this there must be an interpretation of the syntax and the formal semantics which readily responds to intuitive deontic notions. Unless this is done we cannot even decide with confidence if the theorems are in fact true.

The third and fourth chapters are included primarily for historical interest. Chapter III describes von Wright's deontic logic and shows that it fails, for several reasons, to meet the Criteria of Adequacy. The fourth chapter presents the work of A. R. Anderson which also fails the Criteria of Adequacy. Out of this rubble, Chapter V attempts to make a fresh start by generating three deontic logics which are simply modifications of the alethic logics described in Chapter II. The semantics of Chapter II which speaks of "possible worlds" is appropriately modified to produce the notion of an "ideal world". However, the paradoxes, particularly Ross's paradox, suggest that ideal worlds are anything but. Here is deontic logic's darkest hour. Rescher and von Wright man the life boats[^1].

The next chapter, Chapter VI, examines Hintikka's contention that quantifiers are necessary. But a preliminary examination indicates that Ross's paradox has not yet been soothed into submission. Furthermore, Hintikka has troubles of his own. At this point I suggest a way out. The word 'ought', it must be noted, is used in normative discourse in at least two ways: to evaluate and to prescribe. There ought to

[^1]: Rescher [1958] and von Wright [1964].
be a street light here,' is an example of the evaluative function of 'ought', while 'You ought to keep your promises,' is an instance of the prescriptive function. Hintikka's valuable work can be salvaged by interpreting his deontic logic as formalizing evaluative uses of 'ought'. On this interpretation the paradoxes dissolve. Chapter VI closes with a discussion of Searle's well-known paper, "How to Derive Ought From Is" and the question of whether ought implies can.

Chapter VII looks at Castaneda's deontic logic because he specifically defends it as a formalization of prescriptive oughts. Castaneda's major contribution to deontic logic is the introduction of the notion of agency as a fundamental element of prescriptives and imperatives. However, I argue in this chapter that his deontic logic fails to adequately capture the logical features of this important concept.

Chapters VIII and IX attempt to construct an improved deontic logic which avoids the difficulties of classical deontic logics. Chapter VIII leads off with the development of a logic which captures the logical features of agency. Since I agree with Castaneda that agency is an essential component of imperatives and prescriptive uses of 'ought', Chapter VIII concludes with an imperative logic, while Chapter IX presents an improved deontic logic based on the foundations built in Chapter VIII. The improved deontic logic formalizes evaluative and prescriptive uses of 'ought' as well as a number of other important ethical distinctions. Chapter IX concludes with a discussion of Rawls' important distinction between summary rules and rules of practice, and
by noting an important limitation of the improved deontic logic.
The net result, I believe, is that monadic deontic logic is alive and quite well.
CHAPTER II
ALETHIC MODAL LOGIC AND POSSIBLE WORLD SEMANTICS

2.1 An Introduction

Von Wright originally viewed deontic logic as a species of modal logic, and the locutions 'It ought to be the case that...' and 'It is obligatory that...' are similar to the locutions in which other propositional modal notions are often embedded ('It is possibly the case that...' and 'It is known that...'), so it should come as no surprise that much of the work in deontic logic has borrowed material from traditional alethic modal logic. The formal study of alethic modal logic has been rather intense since it was introduced by C. I. Lewis¹. To fully appreciate some of this work and its influence on deontic logic, there are two important things to have: a basic, philosophically relevant interpretation of some standard alethic logics, and a decision procedure for those logics. The procedure introduced below makes those interpretations perspicuous and can easily be altered to provide decision procedures for a variety of extant deontic logics.

2.2 Standard Alethic Logics

Three well-known alethic logics are T, S₄, and S₅. Briefly, they can be syntactically developed as follows. To any complete and

¹Lewis [1918] and Lewis and Langford [1932].
consistent basis for the propositional calculus, add

(A1) \( Lp + p \)

(A2) \( L(p + q) + (Lp + Lq) \)

as axioms, and the following rule of inference:

(R1) If \( A \) is a theorem, then \( LA \) is a theorem.

The resulting logistic system is known as \( T \) and contains the following among its derived rules:

(DR1) If \( A \rightarrow B \) is a theorem, then \( LA \rightarrow LB \) is derivable as a theorem.

(DR2) If \( A \rightarrow B \) is a theorem, then \( MA \rightarrow MB \) is derivable as a theorem.

Some interesting theorems are:

(T1) \( p + Mp \)

(T2) \( L(p \& q) \equiv (Lp \& Lq) \)

(T3) \( Lp \equiv \neg Mp \)

(T4) \( M(p \lor q) \equiv (Mp \lor Mq) \)

(T5) \( Lp + L(q + p) \)

(T6) \( L\neg p + L(p \rightarrow q) \)

(T7) \( Lp + Mp \)

(T8) \( L(Lp + p) \).

Except for (T5) and (T6), this list contains expressions that fairly well accord with our intuitions about the notions of necessity and possibility. And there are arguments that (T5) and (T6) are, despite

\(^{2}\) Read 'Lp' as 'p is necessarily true' and 'Mp' as 'p is possibly true'.
appearances, innocuous\(^3\). Another interesting feature of \(T\) is that it contains an infinite number of modalities. A modality is any unbroken sequence of zero or more monadic operators, \(\text{viz.}, '1', 'L',\) and 'M'. 'LLp', for example, is not equivalent to 'Lp'. The other two modal systems we shall mention, \(S_4\) and \(S_5\), have fourteen and six, respectively.

The system \(S_4\) is generated by adding to \(T\) the following axiom:

\[(A3) \quad Lp \rightarrow LLp.\]

A significant difference between \(T\) and \(S_4\) is that the special axiom added to \(S_4\) permits the derivation of theorems which allow the reduction of modalities:

\[(T9) \quad Lp \equiv LLp\]
\[(T10) \quad Mp \equiv MMp\]
\[(T11) \quad LMp \equiv LMLMp\]
\[(T12) \quad MLp \equiv MLMLp.\]

Consequently, every modality in \(S_4\) reduces to either \('p', 'Lp', 'Mp', 'LMp', 'MLp', 'LMLp', 'MLMLp'\) or their negations.

\(S_5\) is generated by adding to \(S_4\)

\[(A4) \quad Mp \rightarrow LMp\]

from which two more reduction theorems may be derived:

\[(T13) \quad Mp \equiv LMP\]
\[(T14) \quad Lp \equiv MLp\]

and every modality in \(S_5\) reduces to either \('p', 'Lp', 'Mp', or their negations.

\(^3\)Hughes and Cresswell [1968], p. 40.
2.3 Truth Trees

The method I will use for interpreting and deciding the validity of modal expressions can be illustrated by interpreting propositional logic and exhibiting a procedure for deciding the validity of expressions of propositional logic. An expression of propositional logic is a tautology, or PL-valid, if and only if its negation has no consistent model. A consistent model of an expression (or a set of expressions) is an assignment of truth values to the propositional variables of that expression (or set of expressions which satisfies the following conditions:

(C.2) Every propositional variable is either true or false, but not both.

(C.-) Any wff, p, is true if and only if \(-p\) is false.

(C.&) \((p \& q)\) is true if \(p\) is true and \(q\) is true, otherwise \((p \& q)\) is false.

(C.v) \((p \lor q)\) is true if either \(p\) is true or \(q\) is true, otherwise \((p \lor q)\) is false.

(C.\(\rightarrow\)) \((p \rightarrow q)\) is true if either \(p\) is false or \(q\) is true, otherwise \((p \rightarrow q)\) is false.

(C.\(\equiv\)) \((p \equiv q)\) is true if either \(p\) is true and \(q\) is true or \(p\) is false and \(q\) is false, otherwise \((p \equiv q)\) is false.

If, for example, the PL expression \('p \rightarrow (q \rightarrow (p \& q))'\) is PL-valid, then its denial has no consistent model. Let us see what happens when we trace the consequences of denying \('p \rightarrow (q \rightarrow (p \& q))'\). If it is false, then according to \((C.\rightarrow)\), \('p'\) and \('- (q \rightarrow (p \& q))'\) are
both true, and if '-(q \rightarrow (p \land q))' is true, then 'q' and '-(p \land q)' are also true. Let us represent this chain of reasoning as follows:

\begin{align*}
1 & \quad \sqrt{-p} \quad \sqrt{(q \lor (p \land q))} \quad \text{Assume} \\
2 & \quad p \\
3 & \quad \sqrt{-q} \quad \sqrt{(p \land q)} \\
4 & \quad q \\
5 & \quad -(p \land q)
\end{align*}

The check marks indicate that all of the consequences of the assumption that the checked expression is true have been added to the list. If the expression on line 5 is true, then according to (C.\&) either 'p' is false or 'q' is false. We consider these alternatives independently of one another by introducing a branch to the list:

\begin{align*}
5 & \quad \sqrt{-p} \\
6 & \quad \sqrt{-q}
\end{align*}

But we cannot, as the left branch requires, consistently suppose that that 'p' is false since, on line 2, we are supposing that 'p' is true. Hence, the left branch violates (C.\&) and it is, therefore, closed.

I shall adopt the convention of placing an 'X' at the bottom of a closed branch. A closed branch does not represent a consistent assignment of truth values to the variables of the expression being tested. The right path in the above tree suffers the same fate.

The branching character of this list has won it the title of truth tree. Truth trees of all sorts are described in Jeffrey [1967]. The rules for checking any type of PL expression are given in Table 1,
and instructions for applying them are given in the flow diagram of Table 2.

2.4 Interpretations of Alethic Logics

The standard interpretation of formulas of alethic logics treats 'Lp' as 'p is true in all possible worlds' and 'Mp' as 'There is a possible world in which p is true'. The interpretations of T, S4, and S5 differ only on what counts as a possible world. These differences will be explained presently.

As before, the negation of a valid expression has no consistent model. For all three of these alethic logics, a model is a set of possible worlds (one of which is also the actual or real world) which are alethic (possible) alternatives to one another according to rules to be presented. Truth values are assigned to all of the formulas of each world in accordance with the PL conditions listed in the previous section and the conditions which follow:

\[(C.L) \quad V(Lp, W_i) = T \text{ if and only if } V(p, W_k) = T \text{ for each } W_k \text{ such that } A(W_k, W_i), \text{ and there is at least one } W_k \text{ such that } A(W_k, W_i)\.
\]

\[(C.M) \quad V(Mp, W_i) = T \text{ if and only if } V(p, W_k) = T \text{ for at least one } W_k \text{ such that } A(W_k, W_i)\.
\]

These rules employ some formal short-hand which should be understood as follows: 'V(f, W_i) = T' means 'The truth value of f in world W_i is true,' and 'A(W_k, W_i)' is to be read as 'World W_k is an alethic alternative to world W_i.' This alternativeness relation can be under-
TABLE 1

\[
\begin{array}{cccc}
\sqrt{A \rightarrow B} & \sqrt{A \land B} & \sqrt{A \lor B} & \sqrt{A \equiv B} \\
\downarrow & \downarrow & \downarrow & \downarrow \\
-A & \land & B & -A & B & -B \\
\end{array}
\]

\[
\begin{array}{cccc}
\sqrt{-(A \rightarrow B)} & \sqrt{-(A \land B)} & \sqrt{-(A \lor B)} & \sqrt{-(A \equiv B)} \\
\downarrow & \downarrow & \downarrow & \downarrow \\
-A & \land & -B & -B & -B & B \\
\end{array}
\]

TABLE 2

List the premises and the denial of the conclusion.

Erase all double negations. Close each path that contains both a sentence and its denial.

Stop. The inference is valid.

Are all paths closed?

NO

Stop. The inference is invalid.

Have all sentences containing '\', or '∧' or '∨' or '≡' been checked?

NO

Choose some such sentence. Check it. Apply the relevant rule of inference to it.
stood in terms of what is conceivable to the inhabitants of some world. If they can conceive of some possible world, then that world is an alethic alternative to their own. Possible worlds which they cannot conceive of are not alethic alternatives to their own world. As an example, we might imagine a possible world in which blue is not a color. Such a world is, perhaps, not one which we can conceive of. Hence, it is not an alethic alternative to our world even though (as we are supposing) it is a possible world. This relation may be either reflexive, transitive, symmetrical, or any combination of the three.

The interpretation of T treats this relation as reflexive. So, in addition to the conditions already listed, a T-model must satisfy one further condition:

(C.Ref) For every \( W_i \), \( A(W_i, W_i) \).

Hence, a world in which 'Lp' is true is an alethic alternative to itself, and according to (C.L) 'p' is also true in that world. Figure 1 illustrates how truth trees are extended into forests which represent T-models and thereby test for T-validity. The forest of Figure 1 represents a model for

(1) \(- (Lp + L(q + p))\).

The only branch in \( W_1 \) is closed which means that \( W_1 \) represents an impossible world. Since \( W_0 \) requires the existence of \( W_1 \), \( W_0 \) is also inconsistent. Hence the denial of (1) is T-valid. The asterisk appended to 'Lp' on line 2 of \( W_0 \) is a reminder that p must be added to any world which is an alethic alternative to \( W_0 \). Both \( W_0 \) and \( W_1 \)
\[ \begin{array}{|c|c|c|}
\hline
\text{W}_0 & \text{FROM} & \text{Assume} \\
\hline
1 \quad \sqrt{-(L_p \rightarrow L(q + p))} & 1 \\
2 \quad \sqrt{L_p \ast} & 1 \\
3 \quad \sqrt{-L(q + p)} & 1 \\
4 \quad \sqrt{L-(q + p)} & 3 \\
5 \quad p & 2 \\
\hline
\end{array} \]

\[ \downarrow \]

\[ \begin{array}{|c|c|c|}
\hline
\text{W}_1 & \text{W}_0-4 & \text{W}_0-2 \\
\hline
1 \quad \sqrt{-(q + p)} & 1 \\
2 \quad p & 1 \\
3 \quad q & 1 \\
4 \quad -p & 1 \\
X & \\
\hline
\end{array} \]

\textit{FIGURE 1}
are alethic alternatives to $W_0$, so 'p' has been added to both. Notice that expressions beginning with 'L-' are replaced with expressions beginning with 'M-'. This is done because they are equivalent and (C.N) can be applied to expressions beginning with 'M'. Similarly, expressions beginning with 'M-' are replaced with expressions beginning with 'L-'.

$W_0$, the arrow, and $W_1$ constitute a path through the forest. A branch, however, does not extend beyond the world in which it occurs. On the basis of the direction of the arrow, $W_1$ is said to be below $W_0$. In a T-forest (representing a T-model) $A(W_k, W_1)$ if and only if $W_k$ is the world immediately below $W_1$ or is $W_1$ itself.

In an S4 model, the relationship between alternatives is both reflexive (as in T-models) and transitive. So, in an S4 forest, $A(W_k, W_1)$ if and only if $W_k$ is anywhere below $W_1$. See Figure 2. The term 'below' is to be understood such that in Figure 2 $W_2$ is below $W_0$. Hence, being an S4 forest, Figure 2 tells us that $W_2$ is an alethic alternative to $W_0$ and since 'Lp' is true in $W_0$, 'p' is true in $W_2$. $W_2$ is an inconsistent world, hence this forest is inconsistent and the model it represents is inconsistent. Hence 'Lp + LLp' is S4-valid. It is not T-valid since in a T forest $W_2$ is not the world immediately below $W_0$ and therefore does not represent an alethic alternative to $W_0$. The new condition which an S4-model must satisfy (in addition to all of the conditions T must satisfy) is

$$(C.Trans) \quad \text{If } A(W_1, W_k) \text{ and } A(W_k, W_j), \text{ then } A(W_1, W_j).$$

An S5-model must satisfy one more condition,
\[
\begin{array}{ccc}
\text{W}_0 & \text{FROM} & \text{Assume} \\
1 & \sqrt{-(Lp + LLp)} & \\
2 & \sqrt{Lp} & 1 \\
3 & \sqrt{-LLp} & 1 \\
4 & \sqrt{M-Lp} & 3 \\
5 & p & 2 \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{W}_1 & \downarrow & \\
1 & \sqrt{-Lp} & \text{W}_0-4 \\
2 & \sqrt{M-p} & 1 \\
3 & p & \text{W}_0-2 \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{W}_2 & \downarrow & \\
1 & -p & \text{W}_1-1 \\
2 & p & \text{W}_0-2 \\
\end{array}
\]

**FIGURE 2**
What this condition means for an S5 forest is that A(W₁,Wₖ) if and only if Wₖ is in the same forest as W₁. Examine Figure 3. Since W₁ is in the same forest as W₂, W₁ is an alethic alternative to W₂, so '¬p' is true in W₁. A word of warning: In the forest of Figure 4 W₀ has two branches each of which requires an alethic alternative. Each branch and the path for that branch must be treated as an alternative forest. Hence, W₁ and W₂ are not both in the same forest and on an S5 interpretation are not alethic alternatives to each other.

The reader may be wondering whether or not all of the tree rules really do capture all of the conditions. The tree procedure is simply a tabular variant of the procedure developed in Hughes and Cresswell [1963]. The same rules are used, I just put the marks on paper in a different arrangement. The justification for this different arrangement is that it is a way of easily extending truth trees to provide an extremely simplified and perspicuous decision procedure. To modify truth trees for alethic logic, the only new tabular notion required is the notion of an alternate tree which corresponds in an obvious way to the notion of an alternate world. So, the adequacy of this method depends on the adequacy of the method of Hughes and Cresswell, and I shall not offer any independent arguments to that effect.

2.5 Deontic Interpretations

The next preliminary is to work out a vocabulary to evaluate deontic logics. As a preliminary to doing that, I wish to point out
\textbf{FIGURE 3}

<table>
<thead>
<tr>
<th>(W_0)</th>
<th>(\text{FROM})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(\sqrt{-(M_p + LM_p)}) (\text{Assume})</td>
</tr>
<tr>
<td>2</td>
<td>(\sqrt{M_p}) (1)</td>
</tr>
<tr>
<td>3</td>
<td>(\sqrt{-LM_p}) (1)</td>
</tr>
<tr>
<td>4</td>
<td>(-p) (W_2-2)</td>
</tr>
<tr>
<td>5</td>
<td>(\sqrt{N-M_p}) (3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(W_1)</th>
<th>1</th>
<th>(p) (W_0-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>(-p) (W_2-2)</td>
<td></td>
</tr>
<tr>
<td>(X)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(W_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>
FIGURE 4
that the meanings could have been assigned to the elements of $T$, $S^4$, and $S^5$ and their models, and a decision procedure developed without once uttering the phrase 'possible worlds'. It could have been done (and has been done) in terms of model sets of wffs or some equally opaque notion, and consistency and completeness could have been proven with respect to whatever was specified as "formula of designated value." Thus, the interpretation of a formal system can be just as philosophically uninteresting and just as formal as the formal system itself. Interpretations begin to become interesting when the objects which constitute the model are given names or described in such a way that one begins to see the model as a way of understanding the world or, what may be more nearly correct, understanding a part of ordinary language.

For example, the syntactic structure (*) composed of '$$', positive integers, '(', and ')' which form well-formed strings if and only if that string is formed in accordance with either

(FR1) A positive integer standing alone is a wff.
(FR2) Two wffs connected with '$$' is a wff.

is of no particular philosophical interest. Being given a list of axioms and rules of inference is of no interest either. Nor is it particularly exciting to know that $V(A) = \emptyset$ if and only if $V(A \times A) = \emptyset$, and that $V(A \times B) = \emptyset$ if $V(A) = \emptyset$ and $V(B) = \emptyset$, otherwise

$V(A \times B) = \emptyset$, and that an expression, 'A', is (*)-valid if and only if $V(A) = \emptyset$ for all $V$-assignments to the variables of A. Nor do philosophers perk up their ears when told that (*) is consistent and complete with respect to (*)-validity. All of this is certainly very
nice—it is in (*)'s favor that it is formally in order—but if told that (*) captures the logic of 'is better confirmed than' or 'intends' or 'is more real than' or 'neither...nor', then our Humean skepticism begins to bubble and counter-examples rattle threateningly.

We do not prize, philosophically, a logic merely because it has a formal semantic basis, consistency proofs, and completeness proofs. We prize it, philosophically, if these conditions are satisfied and if it also tells us something about the logical structure of our discourse. If 'p ∨ q' means '¬p & ¬q', then (*) represents a complete and consistent basis for propositional logic (given an appropriate set of axioms and transformation rules) and formalizes, in an unwieldy way, the logic of the ordinary words 'and', 'or', and 'not', though, as we all know, it does not formalize the logic of all uses of these words. When we are told what is analytically true for a concept, e.g., that '∅' means 'true' and '∅' means 'false', our ability to distinguish valid and invalid arguments which employ those concepts is also increased, and so is our understanding of the concept. I take this to be the paradigm of philosophical progress.

To be of philosophical interest, we need, in addition to a syntax and a semantic model, a syntactic reading and a semantic reading. By 'reading' I mean some method of letting the formulas stand for certain classes of English (in my case) sentences. For example: 'p ∨ q' means 'neither p nor q'. On this syntactic reading, (*)& can be seen as attempting to capture the logical behavior of a piece of ordinary discourse. Giving what I am calling a syntactic reading is
often referred to as an intuitive basis or an *informal interpretation* of the calculus. A semantic reading tells us how to read the elements of the formal semantics in English. I cannot call this a formal interpretation. Common usage has pre-empted this phrase to speak of the purely formal relationship between a syntax and its formal semantics. I shall refer to the relationship between the formal semantics and English as a *valuation interpretation*. An example of a semantic reading for (\(A\)) has already been suggested, i.e., \(V(A) = 0\) means 'A is true'. A valuation interpretation tells us how to read all of the semantic elements. An informal interpretation tells us how to read all of the syntactic elements. If a logic (a syntactic structure) has been given an informal interpretation, a formal interpretation (a formal semantics) and a valuation interpretation, then it is a candidate for being a fully interpreted logic. I say a 'candidate' because something else is required. There must also be some intuitive glue between the informal interpretation and the valuation interpretation so that the valuation interpretation can be appealed to to refine our understanding of what can and what cannot be an English reading of wffs using the concepts which the informal interpretation impresses upon the syntactic elements. Granted, the kind of English sentences associated with the semantic elements has a huge influence on the strength of the intuitive connection. But if we had interpreted the semantic structures for \(T\), \(S4\), and \(S5\) in terms of just trees and forests, or ordered \(n\)-tuples and designated values, we would not have as much of an understanding of the notions of possibility and necessity as I think we do when those same structures are interpreted in
terms of possible worlds.

The degree to which a logic is fully interpreted depends in large measure on how much the informal interpretation and the valuation interpretation have in common. If they have a large number of common notions, the stronger the glue will be. For example, the syntactic reading 'p is possible' and the semantic reading 'p is true in some possible world' are closely related. Once such an intuitively satisfying relationship is found, a logic is more or less fully interpreted depending on the strength of the intuitions.

Deciding whether or not a fully interpreted logic is adequate is not an easy matter. This task is made more or less easier as the logic is more or less fully interpreted. But a Criterion of Adequacy is not difficult to specify: A fully interpreted logic is adequate if and only if, given that interpretation, all of the valid formulas are true. It is not always an easy matter to determine whether or not the Criterion has been satisfied. Consider two possibilities. First, given a syntactic reading for a formula, that formula may be quite obviously and irremediably false despite the fact that there is a semantic model on which that formula is valid. For example, 'Lp \rightarrow p' is a deducible theorem of S4 and is also a logical truth given Hintikka's model sets semantics for S4. Yet if we read 'Lp' as 'p ought to be true', then 'Lp \rightarrow p' seems quite false and there is no apparent way to render the relationship between the theorem, the model and the reading as an adequate one. But it is possible that a completely different informal interpretation of the
formulas might produce an adequate full interpretation. A prima
facie reasonable alternative in this instance permits only impera-
tives as values of the variables. A deontic logic of this type is
considered in Chapter VII.

Secondly, given a syntactic reading for a formula, that form-
ula may appear to be false despite the fact that the relationship
between the object language, semantic model, syntactic reading and
valuation interpretation seems quite adequate in most other respects.
What I have in mind are the paradoxes of material implication. On the
surface, \(-p \rightarrow (p \rightarrow q)\) seems false especially when read as 'A false
proposition implies any proposition whatever.' In this case, the
adequacy of the propositional calculus can be salvaged by appealing
to the valuation interpretation (provided it is stated in terms of
truth and falsity) to see that \(-p \rightarrow (p \rightarrow q)\) should not be informally
interpreted as saying 'A false proposition implies any proposition
whatever.'\(^4\) This second case illustrates the fact that to a certain
extent the valuation interpretation can prescribe syntactic readings.
The extent to which a valuation interpretation should be extended
legislative powers will no doubt always be arguable, but doing so
invariably restricts the applicability of a logic. For example, the
list of 'if...then' locutions which the propositional calculus cannot
handle grows constantly.

I intend to point to examples of both of the possibilities
listed in the previous paragraph; to argue that some deontic logics

\(^4\) For an effective argument of this see Kielkopf [1972].
(the sanction based deontic logics of A. R. Anderson) are inadequate because they admit of false, vague, or senseless theorems on their full interpretations; to argue that another class of deontic logics (the standard deontic logics other than Anderson's) long thought ravaged by paradoxes, can be given new life and succor by developing a clear valuation interpretation and appealing to that interpretation for a more sober reading of the putatively paradoxical theorems. And, finally, to show as a consequence of this wringing out process, that the standard deontic logics are only adequate to deal with one of several clearly distinguishable senses of 'ought'.

In connection with the last point, I am not going to say very much about legal obligation. Perhaps the results of the present work can be extended to cover legal uses of 'ought', but, perhaps not. I am not familiar enough with legal theory to say one way or the other. But this will not keep me from appealing to legal examples on occasion.

2.6 When Is a Logic Deontic?

A Logic is deontic if and only if there is a syntactic element (not necessarily primitive) designed to formalize, in some way, meaningful uses of the word 'ought' in normative discourse. And if there is a semantic specification of what a valid formula is, then transformation rules are unnecessary. Hintikka's deontic logic is an example of this (as are his alethic and epistemic logics). Deontic logic, as I see it, is quite simply the logic of 'ought'. Whether a
logic is deontic or not is thus dependent upon its designer's intentions or the intentions of anyone reinterpreting a logic which was originally intended to do something else. One might (informally) interpret \('Lp' in S4, to repeat an example, to read 'It ought to be that p'. This reinterpreted S4 is a deontic logic, but, as we have seen, if we assume that \('p' is a proposition, then S4 does not have an adequate full deontic interpretation. Hence, there is no reason to believe S4 is an acceptable deontic logic.

There are, then, any number of deontic logics and the present work will examine several of them. How many deontic logics satisfy the Criterion of Adequacy? Is there only one correct deontic logic that deontic logicians are attempting to discover, or are there several? And if there are several, how is that to be explained? One possibility is that two deontic logics are both adequate because they turn out to be deductively equivalent which means that whatever theorems are found in one are also in the other. Secondly, there could be two adequate deontic logics which are not deductively equivalent if the theorems of one are contained in the other. They would have to be complete with respect to different (but related) semantic bases. Of these two logics, the one which is contained in the other would only be adequate to deal with a semantically distinguishable sub-class of locutions dealt with by the containing logic. For example, T contains propositional logic. Propositional logic and T can both be interpreted adequately--they simply perform different analytic tasks.
There is one last alternative—two deontic logics such that neither contains the other. Is it possible that they are both adequate? I believe so. A major thesis of the present work is that the word 'ought' is used in a variety of ways in ordinary discourse and some of them are sufficiently different to warrant different logical treatments. There is no a priori reason for supposing that these various logical senses of 'ought' have a common core which can be captured by a single deontic logic. If anything, subsequent chapters will show that this is false.

What accounts for the various senses of 'ought'? Is there a hedonistic sense, a utilitarian sense, an intuitionistic sense? After all, we do speak of what a utilitarian "means" by ought. Yes, unfortunately, we do, but there are different senses of 'senses'. The utilitarian sense of 'ought' tells us the necessary and sufficient truth conditions for sentences which employ 'ought'. It is not, as a deontic logic is, a specification of what can consistently be said with sentences which contain 'ought'. A set of truth conditions must be in accordance with the consistency conditions for a given use of a word. Even an apparent relativist such as C. L. Stevenson is prepared to reject a theory which requires a single action to be both right and wrong\(^5\). Either the consistency conditions are independent of truth conditions or communication using the word in question is impossible. Otherwise if a utilitarian and an intuitionist "disagree" over whether they ought to torture one man to eliminate poverty and disease, it is

\(^5\)Stevenson [1942], p. 73.
only in a metaphorical sense of 'disagree' since they are not talking about the same thing. One is saying 'Op' and the other is saying '¬Op' which might look like a contradiction, but it is not since the two occurrences of '0' have different meanings. The independence of consistency conditions from truth conditions seems to be a necessary condition for communication and rational discourse.

However, it has been argued that utilitarianism is incompatible with any deontic logic which contains

\[(2) \quad O(p \land q) \equiv (Op \& Oq)\]

as a theorem\(^6\). It is a theorem of all standard deontic logics. (2) may be read as 'it is obligatory that both p and q if and only if it is obligatory that p and it is obligatory that q.' A utilitarian, at least an act utilitarian, defines obligation in terms of the best possible action given the circumstances: Do that action which leads to the most good. Consequently

\[(3) \quad Bp \ (\text{It is best that } p) \equiv Op.\]

Therefore,

\[(4) \quad B(p \land q) \equiv O(p \land q).\]

But if

\[(5) \quad O(p \land q) \equiv (Op \& Oq)\]

and if

\[(6) \quad (Op \& Oq) \rightarrow Op\]

and if

\[(7) \quad Op \equiv Bp,\]

\(^6\)Castaneda [1968].
then

\[(8) \quad B(p \& q) \rightarrow Bp.\]

But if we assume that 'p & q' describes the best course of action, it cannot follow that 'p' is also the best possible action. Thus utilitarians cannot avail themselves of the standard deontic logics and the standard deontic logics do not adequately reflect the utilitarian sense of 'ought'.

This argument is effective against any normative theory which defines 'ought' in terms of a superlative. A masochistic theory, for example, might define 'ought' in terms of the worst possible course of action. (7) is paradoxical for this theory as well. Definist theories are currently on the wane and more attention is being paid to the logic of moral utterances; moral philosophers are asking what sort of moral claims one can consistently conjoin. But even if definism is not currently de rigueur among ethical theorists, I prefer to not adopt a deontic logic which begs the question against them. I am attempting, among other things, to look for a neutral deontic logic; one that is consistent with as many reasonable normative theories as possible. If the standard deontic logics are inconsistent with utilitarianism, then the standard deontic logics are not the deontic logician's stone for which I am searching. However, according to the introduction, I plan to defend the standard deontic logics as adequate formalizations of one sense of 'ought'. Clearly, I must disarm Castaneda's argument.
Consider some examples of \( O(p \& q) \):

(9) John ought to keep his promises, and, I almost forgot, return the rifle.

(10) It ought to be that John keeps his promises and Morton saves his money.

(11) Everyone ought to refrain from killing and stealing.

In all three cases it seems perfectly reasonable to accept \( O(p \& Oq) \) as a consequence. In fact it seems intuitively unreasonable to not accept it as a consequence. The reason being that 'p' and 'q' are independent of one another; 'p' is obligatory in one situation and 'q' is obligatory in another. These actions are not morally related to one another and suggest that the interpretation of 'Bp' should be modified to read 'It is best that p in this situation.' Hence, the utilitarian may hold (5) as well as (8) without paradox. But it must be understood that if 'p' and 'q' are the best courses of action under the circumstances, then 'p' and 'q' must be relevant to different sets of circumstances. In other words, it is questionable whether \( O(p \& q) \) is the proper translation of

(12) You ought to tear up this letter and throw it away.

Suppose you have written a letter which if sent will do irreparable damage to the lives of several people. Suppose further that for this reason I have convinced you of (12). You say, "Yes, you have convinced me that I ought to tear up this letter and throw it away. But, tell me something, ought I tear up the letter?" You have missed an important point underlying the conversation, and it is a logical point.
But, suppose you ask "Ought I throw it away?" A reasonable reply might be that it will do no good just to throw the letter away—it must also be torn up. So here is a case where 'p' and 'q' are relevant to the same situation, hence (12) apparently cannot be translated as '0(p & q)'. However, I believe it can. We may reasonably suppose that 'p' and 'q' are performed in different situations. What does 'it' refer to in (12), the whole letter or the torn up letter? If the letter is torn up, then we have a different situation from that in which the letter is whole, and the obligation to throw it away is relevant to the situation in which the letter is torn up. Apparently, 'and' may be used to express a sequential set of obligations, for example, "You ought to go back there, clean up the mess you made, and apologize." Here we have a set of obligations in which one obligation is relevant to the situation resulting from the satisfaction of another obligation.

A question which arises here is what makes a situation relevant to one obligation and not another. This is one of many tasks confronting any ethical theorist. There seems to be an intuitive difference here, and it might be worthwhile to find a theoretical foundation for it. Insofar as this is a task of any ethical theorist, it is also a task of the utilitarian. There seems to be an intuitive difference between (10), (11), and (12), on the one hand, and (13) You ought to keep your feet on the ground and your eye on the ball, on the other. Here is a case in which the two actions relate to the
same situation and cannot be construed as sequential in the manner of (12). It would be unreasonable to infer that it is best to keep your feet on the ground in this situation for a utilitarian since (13) is saying that something else is best. We must conclude that '0(p & q)' is an incorrect translation of (13). I hope that it will become clear in the sequel that the logical structure of (13) is much more complex than '0(p & q)' suggests. Notice that 'p' and 'q' represent two different actions when '0(p & q)' is used to translate (10), for example, but they represent two different characteristics of the same action when it is used to represent (13). It is simply asking too much of '0(p & q)' to expect it to play both roles. In order to express this difference we must resort to quantifiers which are introduced in Chapter VI.

This conclusion should be just as welcome to a non-utilitarian since the distinction it is based on (two actions vs. two properties of a single action) is not peculiar to utilitarianism. The proper treatment of this distinction must be delayed until the last chapter when a full complement of formal devices will be available. Interestingly enough, in Chapter IX I will suggest that there is a separate utilitarian sense of 'ought' which is not captured by the standard deontic logics. My reasons, however, will be quite different from Castaneda's. It does not follow, to repeat the point of this discussion, that the various senses of 'ought' are a function of one's particular moral views. Instead, the need for a utilitarian sense of 'ought' argued for later suggests to me that utilitarians and non-utili-
Itarians have seized on different senses of 'ought' as their paradigms of what 'ought' means and have erroneously argued that every use must coincide with that paradigm. And that is surely a mistake.
3.1 Introduction

Historically, the first deontic logic to appear in English was that of G. H. von Wright\(^1\). His work was preceded by that of two others, both in German, neither of which has gained much attention\(^2\). This lack of notoriety probably stems from the fact that both of these deontic logics contained theorems which were almost universally felt to be counter-intuitive. For example, Mally’s deontic logic contains the theorem 'Op = p' which even Mally regarded as strange. But he argues that the presystematic notion of ought is ambiguous and his logic helps us to see this. Surely, the common sense of ought is not so ambiguous that we fail to see that all obligations are fulfilled—they notoriously are not. Grelling avoids this absurdity but not others. His logic contains '(p & O-p) → Op' as a theorem which apparently says that if you do what you are forbidden to do, then you ought to do it; a principle often appealed to by government officials, but a principle which

\(^1\)von Wright [1951a] and von Wright [1951b].

\(^2\)Mally [1926] and Grelling [1939]. For a longer discussion of these logics see Hilpinen and Føllesdal [1971].
is unacceptable.

For the purpose of examining contemporary work in deontic logic, von Wright's 1951 entry serves as an appropriate introduction. It is a simple logic, indeed, it is too simple. But despite that simplicity, it might be looked upon as the deontic core of other deontic logics that go beyond it. In fact, deontic logics which contain von Wright's deontic logic are now called standard deontic logics. I will follow this usage. So, an examination of it will plunge us directly into the formalism of deontic concepts without having to concern ourselves with such interesting but complex considerations as deontic operators combined with alethic modalities and quantification.

3.2 The Formal Structure of \( P \).

Its formal structure is as follows:

First, the formal language of \( P \) has the following syntactical elements:

\( (E1) \) Variables: \( p, q, r, \ldots \), with or without subscripts, an unlimited supply, representing the names of act-types.

\( (E2) \) The usual truth-connectives: \( - \), \&, \textcolor{red}{\large\textbf{v}}, +, \equiv.

\( (E3) \) Performance-connectives: \( - \), \&, \textcolor{red}{\large\textbf{v}}, +, \equiv, \) which are strictly analogous to the truth-connectives of negation, conjunction, disjunction, material implication and material equivalence, respectively.

\( (E4) \) Three deontic operators: \( P, O, \text{ and } F \).

\( (E5) \) Punctuation marks: ( and ).

The reader may have noticed that it is difficult to distinguish the

\textsuperscript{3}Hilpinen and Føllesdal [1971], p. 13.
symbols for performance-connectives from those for the truth-connectives. This produces no difficulties, since, as we shall see in the formation rules, performance-connectives connect names of acts and truth-connectives connect P-sentences. Since no ambiguities can arise, the complexities of using distinguishable symbols can be avoided. The well-formed formulas (wffs) of P are called P-sentences and a formula is a P-sentence if and only if it is formed in accordance with the following rules:

(FR1) If A is a wff of propositional logic, then P(A) is a P-sentence.

(FR2) Truth functions of P-sentences are P-sentences.

Note, however, that connectives which occur within the scope of 'P' are performance-connectives because they connect variables which are the names of acts. Connectives which do not occur within the scope of the operator 'P' connect P-sentences. P-sentences are either true or false (but not both), hence these connectives are truth-connectives.

The other deontic operators are defined as follows:

(Def.P) 'O' is short for '¬P-'.

(Def.F) 'F' is short for 'O-'.

The resulting syntactic structure is common to all standard deontic logics. The major difference is that most others interpret the variables as propositions and therefore do not distinguish between performance-connectives and truth-connectives. Von Wright subsequently adopted this interpretation.

\[\text{von Wright [1968], pp. 14 and 16.}\]
Von Wright has also questioned (Def.P). A list of all possible actions cannot be made, he argues, because technological advances will undoubtedly open up entirely new avenues of human behavior (snowmobiling is a good recent example). He says, "As new kinds of act originate, the authorities or norms may feel a need for considering whether to order or to permit or to prohibit them to subjects." Hence, all actions can be divided into two groups: Those subject to norm by the authorities and are, therefore, required, permitted, or forbidden, and those that are not. It follows that

\[ (1) \quad \neg p \rightarrow p \]

is false if 'p' is an act which is not subject to norm by the norm authority. We cannot reject (1) and maintain (Def.P).

I see no reason for rejecting (1). There is, of course, a sense of permission for which (1) does not hold, but it does not strike me as being a moral sense of permission. That sense of permission (von Wright calls it the "strong" sense) requires a norm authority (which might take any number of forms—the Bible, God, or the Sovereign) and it is not at all clear to me that the moral sense of permission is of this sort, that is to say, the basis of moral permissions need not be a norm authority which has not cataloged every act as obligatory, permitted, or forbidden. On the contrary, I am inclined to think that the moral quality of an action is determined not by a norm authority but by the characteristics of the action (which may or may not include its consequences). We may hope that a norm authority would take these mat-

---

\(^5\) Von Wright [1963], p. 86.
ters into consideration in making its decision. Since it might not, we have a natural tendency to distinguish between what authorities permit and what morality permits, and, it seems, morality permits what it does not forbid.

Returning to the formation rules, they specify that

\[\begin{align*}
(2) & \quad Pp \\
(3) & \quad 0(p \lor q) \\
(4) & \quad Op + P(p \& q)
\end{align*}\]

are \(P\)-sentences, and that

\[\begin{align*}
(5) & \quad P0p \\
(6) & \quad p \lor 0q \\
(7) & \quad p \lor q
\end{align*}\]

are not \(P\)-sentences. It should be noted that an expression is well-formed only if all variables which occur in that expression fall within the scope of a deontic operator.

3.3 The Interpretation of \(P\)

\(P\) was originally touted as a calculus of act names; that is, a \(P\)-sentence is an act name or a molecular complex of act names preceded by an operator. In response to the question "What are the 'things' which are pronounced obligatory, permitted or forbidden, etc.?" von Wright answered with confident simplicity, "We call these 'things' acts." So von Wright takes the variables to be act names as expressed by the infinitive of a verb. Then '\(p \& q\)', for example, is read as 'to

\[^6\]von Wright [1951b], p. 2.
go and to stay'. The latter cannot stand alone in English, so the former is not well-formed in $P$. According to the rules, 'p & q' may be transformed into a wff by placing it within the scope of a deontic operator. Hence 'P(p & q)' is well-formed in $P$ and its counter-part in English is 'It is permitted to go and to stay,' which is surely an English sentence. The other two deontic operators are to be read as 'It is obligatory...' and 'It is forbidden...', respectively. As an aid in reading wffs of $P$, the following list is provided:

(8) Fp (It is forbidden to go),
(9) O p v P q (It is obligatory to go or it is permitted to stay),
(10) P(p v q) (It is permitted either to go or to stay).

Not all readings are unproblematic. Consider

(11) 0(p + q).

The most plausible rendering seems to be 'It is obligatory to stay if to go', which does not make sense in English. Von Wright offers an alternative interpretation: 'Doing p commits one to do q.' This reading is unacceptable. According to procedures to be presented in the next section,

(12) F p + 0(p + q)

is a theorem of $P$. It has been effectively criticized by R. M. Chisholm. If the consequent is read as von Wright suggests, (12) states that doing a forbidden action commits us to doing any action. On the contrary, moral intuitions seem to suggest that if we are com-

\footnote{Chisholm [1963].}
mitted to anything at all on the basis of doing a forbidden act, we are committed to somehow righting the wrong that we have done, certainly not to do any action at all. This particular difficulty of \( P \) is now known as the paradox of derived obligation.

Another difficulty with von Wright's interpretation of \( P \) concerns his distinction between act-types and act-individuals. Stealing is an act-type. The robbery of the corner grocery store last night is an act-individual. By an admittedly arbitrary choice, von Wright selects act-types as the values of the variables of \( P \). Luck was against him. The less important reason against that choice is that it may beg the question against act theories of obligation. According to an act theorist it is individual acts that are known directly to be good or bad, right or wrong. For example, Rawls distinguishes between act utilitarianism and rule utilitarianism as follows \(^8\). The act utilitarian justifies individual actions on utilitarian grounds and moral principles are mere summaries of past individual judgments. The rule utilitarian justifies individual actions on the basis of their falling under a rule. It is the rules that are justified by utilitarian considerations. Hence, a rule utilitarian would find \( P \) to be a cordial formalization of moral claims since '\( Fp \)' must be read as 'Stealing is forbidden', i.e., '\( Fp \)' means that every instance of stealing is forbidden. This means that \( P \)-sentences can only be used to espouse moral principles, never, except by implication, to express what some person must do in this particular case. Those moral statements which an act theorist feels

\(^8\)Rawls [1955].
most confident in making cannot be expressed in \( P \). For him, general moral principles are empirical generalizations and are justified inductively. The general claim 'It is forbidden to steal' is justified when it is discovered the lion's share of cases of stealing also happen to be actions which are wrong. Thus, for an act theorist, act individuals are important components of moral reasoning.

The applicability of \( P \) could be greatly increased by changing the arguments of deontic operators to functions of act-individuals and quantifying over them. A wff of this revised version of \( P \) might be something like \((x) P(Lx)\) which would be read 'Every instance of act-type 'L' is permitted.' The quantified version of \( P \) could handle such inferences as 'Stealing is forbidden, and George's taking the watch from the locker is stealing; therefore, George's taking the watch from the locker is forbidden.' But as originally interpreted by von Wright, \( P \) is an inadequate basis for either an inductive deontic logic or a quantified deontic logic.

Next, more difficulties arise when we consider what it means for an act-type to have a performance-value. It is in virtue of performance-values that we may determine whether or not we have fulfilled an obligation, violated a prohibition, or whether doing \( 'p \lor q' \) is the same thing as doing \( '-p \rightarrow q' \). Given that \( 'p' \) and \( 'q' \) represent act-types, none of these things may be determined because the claim that act-types have performance-values (are either performed or not) is either false or difficult to make sense of. Can we answer the question: Has the act type of stealing been performed or not? It is difficult to
see that any answer is possible. There have been, of course, instances of stealing in the past, there are probably some occurring at this very moment, and, no doubt, there will be instances in the future. Obviously it is not act-types that have performance-values but act-individuals; an important difference which von Wright senses when he says "The performance or non-performance of a certain act...we shall call performance values." What he fails to recognize is the notion of a performance value is inconsistent with his choice of act-types as the arguments of deontic operators and treating them as being strictly analogous to truth-functions.

Finally, how are we to interpret \( P(p \& q) \)? There are two reasonable possibilities, (i) All actions which are of act-type 'p' and which are also of act-type 'q' are permitted, and (ii) All actions of act-type 'p' and all actions of act-type 'q' are permitted. Either choice is disastrous. If we choose (i), then

\[(13) \quad P(p \& q) \rightarrow (Pp \& Pq)\]

which is a theorem of \( P \) is false since the antecedent says only that actions which have both characteristics are permitted. If we choose (ii), then

\[(14) \quad (Pp \& Pq) \rightarrow P(p \& q)\]

which is not a theorem of \( P \) appears to be a deontic tautology. In Chapter VI it will become clear that all of these interpretational difficulties are avoided when the variables range over act-individuals instead of act-types.

---

\(^9\text{von Wright [1951b], p. 2. Emphasis added.}\)
3.4 The Decision Procedure for $P$

Von Wright's procedure for distinguishing between $P$-sentences which are deontically valid and those which are not relies upon truth tables and the notion of a $P$-constituent. The $P$-constituents of a $P$-sentence are found in the following way. First, replace all occurrences of '0' with '-P-'. Next, replace the expression within the scope of each 'P' operator with its disjunctive Boolean normal form. Lastly, distribute the 'P' operator to each disjunct. Each resulting disjunct is a $P$-constituent of the original $P$-sentence. As an example, consider the sentence

\[(15) \quad P(p \land q).\]

The expression which falls within the scope of 'P' is

\[(16) \quad p \lor q\]

and its disjunctive Boolean normal form is

\[(17) \quad (p \land q) \lor (-p \land q) \lor (-p \land -q).\]

Replacing (16) with (17) in (15) yields

\[(18) \quad P(p \land q) \lor P(-p \land q) \lor P(-p \land -q)\]

which indicates that (15) has three $P$-constituents.

The $P$-constituents of any $P$-sentence containing $n$ act names are found as follows: For $n$ act names $(p_1, \ldots, p_n)$ there are $2^n$ different conjunction names of acts which are possible constituents of a disjunctive Boolean normal form of any molecular complex of $n$ act names. Each

---

10 In subsequent discussions, von Wright offers an axiomatization of $P$. But, strangely, he continues to define theorems of $P$ in terms of truth table tautologies. For this and the further reason that the truth table technique better exhibits the peculiarities of $P$ I shall ignore its axiomatic development. For details see von Wright [1967], p. 136 and von Wright [1968], pp. 7 and 20.
possible conjunction name when preceded by the operator 'P' is a deontic unit of the deontic realm of the acts named by \( p_1, \ldots, p_n \). The deontic realm itself is the disjunction of all of the deontic units. The difference between a P-constituent and a deontic unit is this. A P-sentence is a truth function of its P-constituents, but the P-constituents for a given P-sentence might not be all of the deontic units under consideration. For example, the P-constituents of (15) are selected from the deontic realm for the acts 'p' and 'q'. There is a deontic unit (a possible P-constituent) which is not a P-constituent of (15), viz., ‘P(p & -q)’.

As an example, consider the deontic realm of the acts named by 'p' and 'q'. The deontic units will number \( 2^2 \) and are ‘P(p & q)’, ‘P(-p & q)’, ‘P(p & -q)’, and ‘P(-p & -q)’. The deontic units are the elements necessary for constructing a truth table for all P-sentences having 'p' and 'q' as constituent act names. In Table 3, the four deontic units of the deontic realm of the acts named by 'p' and 'q' are to the left of the double line and all possible combinations of their truth values are there exhausted with one exception—the table does not include the case when all of the deontic units are false. This idiosyncrasy is the result of the Principle of Permission which is the subject of the next section. The truth values of the expressions which are to the right of the double line are functions of the deontic units (on the left of the double line) as will now be illustrated using

\[ (20) \quad P(p \equiv q) \]
as an example. Since the disjunctive Boolean normal form of \( p \equiv q \) is
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(p \land q)$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$P(p \land \neg q)$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$P(\neg p \land q)$</td>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$P(\neg p \land \neg q)$</td>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$p$</td>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$\neg p$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$P(p \lor q)$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
</tr>
<tr>
<td>$P(p \rightarrow q)$</td>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$P(p \equiv q)$</td>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
</tbody>
</table>

TABLE 3
'(p & q) v (-p & -q)', we know that (20) is a truth function of the P-constituents 'P(p & q)' and 'P(-p & -q)'. Note that the P-constituents of (20) are only two of the four deontic units of the deontic realm of the acts named by 'p' and 'q'. On any line of the truth table where both of these deontic units are false (The P-constituents of (20)), then (20) will be false, otherwise it will be true. Thus, on any line of Table 3 where columns (1) and (4) are both false, then there will also be an 'F' on that line under column (10). Otherwise the value of column (10) will be true.

We are now in a position to see how one may determine whether any given P-sentence is a theorem of P or not. Let

\[(21) \quad (Pp \& 0(p + q)) + Pq\]

which corresponds to the argument

\[(22) \quad 0(p + q)\]

\[Pp\]

Therefore: Pq

serve as a sample test sentence. The procedure depends upon reducing the test sentence to a truth function of its P-constituents. In this particular test sentence there are two distinct act names. Hence the P-constituents of the test sentence will be drawn from the deontic units of the deontic realm of the acts named by 'p' and 'q'.

The antecedent of (21) is the conjunction of Pp and 0(p + q). We must reduce both of these conjunctions to a truth function of their P-constituents. Since the disjunctive Boolean normal form of 'p' (in terms of 'p' and 'q') is '(p & q) v (p & -q)', 'Pp' reduces to
(23) \[ P(p \land q) \lor P(p \land -q). \]

By (Def.0), 'O(p \land q)' is equivalent to '\(-P(p \land q)\)' which is equivalent to

(24) \[ -P(p \land -q) \]

which is simply the negation of one of the deontic units. The consequent of (21) reduces to

(25) \[ P(p \land q) \lor P(-p \land q) \]

in the same manner as 'Pp' described above. So the test sentence is equivalent to

(26) \[ ((P(p \land q) \lor P(p \land -q)) \land -P(p \land -q)) + (P(p \land q) \lor P(-p \land q)) \]

which is a truth function of only three of the four possible deontic units in the deontic realm of the acts named by 'p' and 'q'.

Table 4 is a truth table for (26) in which all possible combinations of truth values of the P-constituents of (26) have been exhausted on the left side of the double line. On the right side the truth values for each connective have been calculated in the order indicated by the numbers above the connectives. Since the main connective is true for all possible combinations of truth values for its P-constituents, it is a deontic tautology. A P-sentence is a theorem of P if and only if it is a deontic tautology. Since (21) is equivalent to (26), (21) is a deontic tautology. Since its corresponding test sentence is a tautology, (22) is a valid argument. We may note that the Principle of Permission places no restriction on Table 4 since (26) is a truth function of only three of the four deontic units of the realm of acts named by 'p' and 'q'. Hence, it is not possible for all of the deontic
\[ ((p \land q) \lor (p \land \neg q)) \land \neg (p \land \neg q)) \Rightarrow (p \land q) \lor (p \land \neg q) \]

<table>
<thead>
<tr>
<th>(P(p \land q))</th>
<th>(P(p \land \neg q))</th>
<th>(P(-p \land q))</th>
<th>((P(p \land q) \lor P(p \land \neg q)) \land \neg (p \land \neg q)) \Rightarrow (P(p \land q) \lor P(-p \land q)))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
<tr>
<td>(T)</td>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
</tr>
<tr>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
<tr>
<td>(T)</td>
<td>(F)</td>
<td>(F)</td>
<td>(T)</td>
</tr>
<tr>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
<tr>
<td>(F)</td>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
</tr>
<tr>
<td>(F)</td>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
<tr>
<td>(F)</td>
<td>(F)</td>
<td>(F)</td>
<td>(T)</td>
</tr>
</tbody>
</table>

TABLE 4
units to be false.

There is a further restriction on \( P \) noted by David Meredith\(^{11}\).
The issue that Meredith raises is one which von Wright did not fully consider in his original paper. That is, how do you treat \( P \)-sentences which contain a null \( P \)-constituent? An example is

\[
(27) \quad \neg P(p \& \neg p).
\]

This is clearly a problem for von Wright since he concludes (intuitively) that expressions such as (27) are contingent. This is embodied in what he calls the Principle of Deontic Contingency. It says, quite simply, that a tautologous act is not necessarily obligatory, and a contradictory act is not necessarily forbidden\(^{12}\). Von Wright admits that his intuitions provide no strong feelings about it, but since he can conceive of no arguments against it, he adopts the principle.

I find that I have no conception of what a tautologous or contradictory act might be. I find it unnatural to apply these terms to anything other than propositions. So let us suppose that the variable in (27) is a proposition, and that (27) is to be understood as saying that it is permitted that a contradiction is true. Its meaning still is not clear, but at least some possibilities come to mind. First, it might mean that it is morally permitted to utter a contradiction. To say that this is a contingency means, I suppose, that it depends on what the contradiction is. But it seems to me that no contradiction may be asserted or defended. To do so is the epitome of irrationality. So if

\(^{11}\) Meredith [1956].

\(^{12}\) von Wright [1951b], p. 11.
there is any moral rule that should be a deontic tautology, it is the
denial of (27). Hence on this interpretation of (27) the Principle
of Deontic Contingency should be abandoned. Secondly, it might be
that in creating the universe God is permitted to create a contradic-
tory state of affairs. Perhaps this is what the Principle of Deontic
Contingency means. Even if we suppose the impossible, that God can do
it, it would probably be best if he did not. It is difficult to imag­
ine what would happen, but Irving Copi provides a reasonable suggest­
on.\textsuperscript{13} What would happen if God created a contradiction? Since any­
thing follows from a contradiction, everything would happen. That does
seem to be a bit much. It seems to me that on this interpretation (27)
should be false. It seems to me that the Principle of Deontic contin­
geney should be rejected.

Meredith draws no deontic conclusions from his purely formal
point. But his result does show that practically nothing is gained by
the Principle of Deontic Contingency. Since the disjunctive normal
form of 'p & \neg p' is empty, the only reasonable way to treat 'P(p & \neg p)'
is as its own P-constituent, and in virtue of the Principle of Deontic
Contingency it may be either true or false. However, Meredith proves
that if the null P-constituent is true, then all other P-constituents
(and all other P-sentences) must also be true.

By the Rule of P-extensionality\textsuperscript{14}

\begin{footnotesize}
\begin{enumerate}
\item Copi [1972], p. 312.
\item von Wright [1967], p. 136.
\end{enumerate}
\end{footnotesize}
(28) PA (where $A$ is any act name) is equivalent to $P(A \lor (p \land \neg p))$

and by the Principle of Deontic Distribution,

(29) $P(A \lor (p \land \neg p))$ is equivalent to $PA \lor P(p \land \neg p)$.

By substitution of equivalents,

(30) $PA$ is equivalent to $PA \lor P(p \land \neg p)$.

Now, suppose it is true that

(31) $P(p \land \neg p)$.

By the rule of addition, (31) yields,

(32) $PA \lor P(p \land \neg p)$.

But by (30) we may replace (32) with

(33) $PA$.

Thus, by conditional proof,

(34) If $P(p \land \neg p)$, then $PA$.

Meredith's restriction, then, is that on any line of the truth table

where the null-constituent is true, every other constituent is also true. Meredith's restriction together with the restriction imposed by

the Principle of Permission eliminates exactly half of the rows of a

truth table which exhausts all possible combinations of truth values of every deontic unit of the deontic realm in question including the

null constituent.

3.5 The Principle of Permission

It is in virtue of the Principle of Permission that Table 1 does not exhaust all possible combinations of truth values of the deontic units. The possibility of all of the deontic units being
false is excluded. This is not an oversight but, rather, the result of what von Wright takes to be the peculiarities of deontic concepts. To see why, let us consider the simplest case; the deontic realm of the single act 'p'. The deontic units are 'Pp' and 'P-p'. To suppose that both units can be false is to admit that

\[(35) \quad -Pp \ & -P-p\]

can be true. Furthermore, (35) is equivalent to

\[(36) \quad -Pp \ & Op\]

and (36), von Wright feels, is clearly counter-intuitive. It reads 'It is not permitted that p (p is forbidden) and it is obligatory that p.' A reasonable person would not insist that some one has an obligation to do something and, at the same time, forbid that person to do it. Furthermore, since whenever we act our act is an instance of any given act-type or not, the falsity of all of the deontic units would mean that we are forbidden to do anything.

It is illuminating to inquire as to what P-sentences are theorems when the restriction that all of the deontic units may not be false is applied but which fail to be theorems when we do allow all of the deontic units to be false. The following table

<table>
<thead>
<tr>
<th>Pp</th>
<th>P-p</th>
<th>P-p v Pp</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

shows that 'P-p v Pp' becomes a theorem when all of the deontic units
are not allowed to be false. If this restriction is lifted, it is no longer a theorem and we have a logic which, with one exception, is no longer deontic but propositional, that is, for each $P$-sentence tautology there will be a corresponding propositional tautology which is obtained by replacing each distinct $P$-constituent with a distinct propositional variable. The exception noted above are those deontic tautologies which are conditionals whose antecedents are the null $P$-constituent and whose consequents are any other $P$-constituent. The proof that without the Principle of Permission $P$ collapses into propositional logic is as follows:

Let $D$ be any $P$-sentence of $n$ act names the $P$-constituents of which are all or some of the deontic units $DU_1, ..., DU_2^n$ but none of which is the null $P$-constituent. Thus, $D$ is some truth function, $f$, of $DU_1, ..., DU_2^n$; that is, the truth value of $D = f(DU_1, ..., DU_2^n)$. Let $S$ be a formula of propositional logic which is obtained by replacing each distinct deontic unit in $f(DU_1, ..., DU_2^n)$ with a distinct propositional variable, $p_1, ..., p_2^n$. Since the function, $f$, remains the same, $f(DU_1, ..., DU_2^n)$ is truth functionally equivalent to $f(p_1, ..., p_2^n)$ and $S$ has the same truth value as $D$. Thus, for any deontic tautology, $D$, there is a corresponding propositional tautology, $S$. But there are no theorems of propositional logic which are analogous to '$Pp \vee P\neg P$'. The propositional analog of this theorem of $P$ is not a theorem of propositional logic. Hence, '$Pp \vee P\neg P$' prevents $P$ from collapsing into propositional logic.
Now let us take up the question of whether the denial of the Principle of Permission is contrary to our logical intuitions or our ethical ones. There seems to be nothing to which we can appeal which clearly shows that a logical intuition has been violated. In fact, if purely logical considerations were at work here, it seems that we would simply fill in the last line of the truth table as a necessary condition of exhausting all of the logical possibilities. The denial of \( 'Pp \lor P-p' \) is (36) and it surely seems unreasonable in some sense. We have already noticed that it is not a logical contradiction. The observation which led von Wright to make the restriction embodied in the Principle of Permission was that "We seem prepared to reject a use of the words, according to which one and the same act could be truly called both obligatory and forbidden."\(^{15}\) Von Wright feels that such a rejection indicates the logical impossibility of all of the deontic units being false. But consider further why we reject such a use.

If an administrator says to a faculty member, "It is your duty to meet your classes when and where scheduled, and, incidentally, you are not permitted to be on campus," he has done something outrageous; but logically outrageous? The administrator is not guilty of uttering a formal contradiction. It is quite possible that an authority create conflicting duties, but it is not possible to create a contradictory state of affairs. The Principle of Permission does not represent a law of logic in the strict sense for the simple reason that it represents a law that can be violated. But neither is it a contingent

\(^{15}\) von Wright [1951b], p. 9.
truth on the order of "Snow is white." What is the logical status of the Principle of Permission?

Anyone may obligate themselves by making promises. If a man promises one day to make a business trip and the next day promises his family he will not leave town, he has obligated himself to leave town and to not leave town. If the man is morally sensitive, it will be quite clear to him that he has contracted both obligations as he anguishes over the decision of which obligation he will obey since he cannot, logically cannot, fulfill both of them. Here, perhaps, is the bedrock of human tragedy, and here is the impossibility that von Wright sensed but failed to identify. Once both obligations have been incurred, it is logically impossible that both of them are fulfilled. But it is not logically impossible to be bound by both obligations. Even so, an enlightened norm authority, an ideal set of norms, or a rational normative theory will not both require and forbid a particular action, and for this reason they will be described by the Principle of Permission. Those unfortunate situations in which we are faced with conflicting obligations I shall call moral dilemmas. They are discussed further in section 9.3. Moral dilemmas are dilemmas precisely because they are morally irrational. The Principle of Permission, on the other hand, is an analytic truth of rational moral discourse.

If it turns out that it is not purely logical considerations that are appealed to to decide the character of deontic logic, then perhaps one might conclude that deontic logic is not properly called "logic". I do not feel this conclusion is warranted. From the previ-
ous considerations it does not follow that a deontic tautology is a statement of the same logical type as

(37) You ought to stay home today.

or

(38) Promises ought to be kept.

Of all of the statements having a deontic content which are logically possible some are deontically necessary, deontically contingent, and deontically impossible. (37), for example, is clearly deontically contingent. Assent to it requires finding out something about the case at hand. Did you promise to stay home? Are there any outstanding overriding obligations? Will anyone be harmed if you stay home? Does your life depend on it? It has been argued that (38) is stronger than (37); that it is analytic. But it is not as clearly deontically analytic as

(39) -(Op & Op-p),

which is an equivalent form of the Principle of Permission. Truths such as (39) are the basic principles of consistency for discourse employing the categories of obligation, permission, and prohibition.

Deontic logic, as I see it, together with alethic logic, epistemic logic, the logic of belief, and other modal logics which have gained popularity in recent years, is more of a logical (in the loose and philosophical sense) enterprise than it is an ethical one. Deontic logic is an attempt to develop a theory which reflects the principles of consistency of normative discourse rather than to discover which moral judgments are true or the more general task of discovering the
truth conditions for particular normative judgments or ethical claims. These tasks belong to ethics. The ultimate task of deontic logic is to produce criteria which we can use to identify consistent and inconsistent normative discourse. These criteria will be above and beyond the principles of logical consistency to which all discourse is subject.

3.6 Some Theorems and Some Limitations

In the previous chapter a Criterion of Adequacy was introduced as an evaluational tool. Reviewing that discussion, an adequate deontic logic is a syntactical structure which

(i) is consistent,
(ii) is complete with respect to a formal semantic structure,
(iii) has an informal interpretation of the syntax,
(iv) has a valuation interpretation of the semantic structure
(v) such that there is a felt intuitive connection between the informal interpretation and the valuation interpretation, an intuitive connection which provides aid in interpreting unclear formulas and testing the truth or falsity of valid formulas.

Von Wright claims, without proof, that (i) and (ii) are satisfied. I will follow this precedent since I am more interested in finding philosophically interesting interpretations of deontic logics than in purely formal considerations. For example, examine the very first theorem listed by von Wright:

16von Wright [1951b], p. 13.
\[(40) \quad (O(p \rightarrow q) \& O\neg p) \rightarrow Oq.\]

It seems fairly obvious that

\[(41) \quad (\neg P(\neg p \& \neg q) \& \neg (\neg P(-p \& -q) \vee P(-p \& -q)) \rightarrow \neg (P(-p \& -q) \vee P(p \& -q))\]

which is equivalent to (40) in virtue of the decision procedure of \(P\) is simply an instance of \('(-p \& \neg (q \vee r)) \rightarrow (r \vee p)\'\) when the deontic units \('P(p \& -q)'\), \('P(-p \& q)'\), and \('P(-p \& -q)'\) are replaced by the propositional variables \('p'\), \('q'\), and \('r'\), respectively. But despite the fact that (40) is a deontic tautology according to von Wright's decision procedure, it is not clear that it is the sort of expression that ought to be a deontic tautology. Assessing the status of (40) is made difficult by the fact that initial intuitions are not clear. In fact, much of deontic logic is plagued with unclear and conflicting intuitions, hence the need for the five criteria listed above. Satisfying the five criteria will help tie down our clear intuitions to a theoretical framework which will help us when our intuitions are not clear and which will show us what we must do to be consistent.

The arguments contained in section 3.3 indicate that (iii) has been only minimally satisfied by \(P\). Indeed, von Wright provides readings for the syntactic elements, but they lead to several anomalies, \textit{viz.}, the notion of an act-type is quite problematic and the reading of \('O(p \rightarrow q)'\) does not seem plausible. But the single most important barrier to a full interpretation of \(P\) is its failure to satisfy condition (iv). \(P\), of course, has been provided with a rather cumbersome semantic structure, but the notions of a \(P\)-constituent and a deontic-
unit which interpret this semantic structure are artificial notions which correspond to no intuitive moral concepts. Hence, they provide no assistance in resolving the problem of whether or not (40) is a theorem. We have also seen that the reading of a deontic unit is incompatible with the specification of logical truth for deontic formulas. Hence, P also fails condition (v) and it is not, as it stands, an adequate deontic logic. This is not to say that this syntactic structure must be abandoned by deontic logic. On the contrary, I believe it to be the foundation of all deontic logic. Eventually I will abandon von Wright's semantics in favor of an alternative world model of the sort described in Chapter II, and I will try different interpretations of those models and the syntax in hope of producing a fully interpreted deontic logic.

Finally, even if we suppose that (iii) and (iv) are satisfied, several of the theorems of P appear to be false given the few guides for interpretation that we have. Those that have been considered paradoxical include

\[(42) \quad 0p + 0(p \lor q)\]

which is known as Ross's paradox,

\[(43) \quad 0-p + 0(p \land q)\]

and

\[(44) \quad 0q + 0(p \land q)\]

which are called the paradoxes of derived obligation, and

\[(45) \quad Fp + F(p \land q)\]
which is known as the paradox of the Good Samaritan. Since, by definition, any standard deontic logic contains \( P \) (and there are several as we shall see), a standard deontic logic will also contain these paradoxical formulas. So a discussion of these paradoxes will appear after a discussion of some other representative deontic logics.

On the basis of considerations presented in this chapter we may look forward to deontic logics which can handle arguments such as

\[
\begin{align*}
\text{(46)} & & 0(p \rightarrow q) \\
p & & \\
\text{Hence: } & & 0q.
\end{align*}
\]

(46) cannot be dealt with by \( P \) because the second premise is not a wff of \( P \). We may also look forward to clear interpretations and quantification.
CHAPTER IV
ANDERSON'S REDUCTION

4.1 Anderson's Proposal

Not long after von Wright's work appeared, Alan Ross Anderson published a major contribution to deontic logic, *The Formal Analysis of Normative Systems*. For a number of years Anderson's logics held the center of the deontic stage and though they have not been discussed much in recent years they represent a novel, but wrong, approach to the logic of obligation.

Briefly, Anderson's work is an attempt to define obligation in terms of the notion of a sanction. Anderson believes that forbidden actions are sanctionable. In fact 'p is forbidden' is defined as 'It is necessary that if p is true, then the sanction is invoked.' Anderson defines deontic operators in terms of alethic ones and in terms of the sanction. This is one reason for having included a brief sampling of traditional alethic modal logic. But even the sanction, Anderson believes, can be reduced to alethic concepts. Hence, he suggests quite strongly that deontic logic can be reduced to alethic logic when he says that "from a formal point of view we may regard deontic logic simply as

---

1Anderson [1956].

2See, for example, Castaneda [1960], Lemmon and Nowell-Smith [1960], Prior [1958], Smiley [1963], and von Wright [1969].
a special branch of alethic modal logic." But this aim requires that we be able to tell when a logic is alethic and when a logic is deontic. I have already suggested that a logic is alethic when there is a syntactical element which is designed to formalize, in some way, meaningful uses of the word 'necessary' (or 'possible'). A deontic logic does the same thing for uses of the word 'ought'. It is simply a matter of how one tries to read the wffs. Anderson's answer to this question is much stronger than mine, and it has the consequence that Castaneda is not really doing deontic logic after all. According to Anderson, a normal deontic logic (i) has propositional logic as a subsystem, (ii) allows the intersubstitutability of provably equivalent expressions, and (iii) has

\[
\begin{align*}
(1) & \quad \Box p \rightarrow Pp \\
(2) & \quad P(p \lor q) \equiv (Pp \lor Pq) \\
(3) & \quad \Box p \equiv \neg Pp
\end{align*}
\]

as theorems, but does not have

\[
\begin{align*}
(4) & \quad Pp \rightarrow p \\
(5) & \quad p \rightarrow Pp \\
(6) & \quad Mp \rightarrow Pp
\end{align*}
\]

as theorems.

According to these criteria, von Wright's $P$ is a normal deontic logic. Every theorem of propositional logic is also a theorem of $P$ provided each propositional variable is replaced with a $P$-sentence. The

\[\text{Anderson [1956], p. 178.}\]
second criterion is identical to von Wright's Principle of \( P \)-extensionality, and (1), (2), and (3) are theorems of \( P \). Castaneda's deontic logic, however, is not a normal deontic logic since \( 'O p + p' \) is a theorem of his deontic logic and \( 'O p + p' \) is equivalent to (5).

The third condition tells us a great deal about Anderson's deontic preferences. He intends to read \( 'O p' \) and \( 'P p' \) (where \( 'p' \) is a proposition) as 'It ought to be that p is true' and 'It is permitted that p is true', respectively. (1) is von Wright's Principle of Permission, (2) is his Principle of Deontic Distribution, and (3) simply reflects the interdefinability of \( 'O' \) and \( 'P' \). (4) is ruled out since

\[
(7) \quad Pp \land P\neg p
\]

is perfectly consistent whereas

\[
(8) \quad p \land \neg p
\]

is not, and (4) would allow us to derive (8) from (7). (5) is ruled out on moral grounds. It is perfectly possible, as well as quite likely, that some forbidden states of affairs are actual. So, from what is actually the case we do not want to be able to derive what is permitted. Let us hope that the world can be better than it is. (6) is undesirable in a deontic logic which combines alethic and deontic modalities. In a normal alethic logic

\[
(9) \quad p \rightarrow M p
\]

would be a theorem, and (9), together with (6), would yield (5) as a theorem which, it has been decided, is undesirable.

I have no argument with Anderson's stipulations given the kind of formal system he is dealing with and providing a reading for. How-
ever, one could provide a perfectly acceptable reading for

\[(10) \quad Op + p\]

which Anderson has ruled out via (3) and (5). Suppose that 'p' stands for an imperative instead of a proposition. On this interpretation (10) may not be false. Indeed, this is the basic insight worked out by Castaneda (Chapter VII) and I prefer a broader characterization of deontic logic (Chapter II). But given Anderson's informal interpretation, his lists of desirable and undesirable theorems are reasonable lists.

A normal alethic logic satisfies the first two conditions of a normal deontic logic and has

\[(11) \quad p + Mp\]
\[(12) \quad M(p \lor q) \equiv (Mp \lor Mq)\]
\[(13) \quad Lp \equiv -M-p\]

as theorems, but fails to have

\[(14) \quad Mp + p\]

as a theorem. It is important that Anderson specifies in an unambiguous way what a deontic logic is and what an alethic logic is to support his claim that he has reduced a normal deontic logic to a normal alethic logic. Next, a description of the reduction.

4.2 Anderson's Reduction

There is, Anderson notes, "an intimate connection between obligations and sanctions." Social groups resort to sanctions or penalties in order to enforce whatever obligations and duties they place upon

---

4Anderson [1956], p. 170.
their members. Think for a moment of the conclusions an anthropologist might draw about the moral code of a society based on certain observations. If he notes that an action results in the application of penalties or sanctions on members of the society responsible for that action, then it seems perfectly reasonable to conclude that actions which do not result in penalties are permitted. But which category is logically basic, prohibition or punishment? It must be Anderson's thesis that punishment is the logically prior category since he defines prohibition and the other deontic properties in terms of it. He is probably wrong. Inflicting misery on someone may or may not be punishment. To decide whether or not it is punishment, we must decide whether or not a wrong has been done. Despite this, Anderson defines the operator 'P' in terms of a sanction as follows:

\[(15) \quad \text{'}Pp' \text{ is short for 'M(p & \neg-S)'}\]

where 'S' is the formal representation of the constant. It is a propositional constant but exactly what proposition is not clear. Anderson aggravates the problem of interpreting 'S' by reading virtually every occurrence of it as 'the sanction'. The major exception to this is an alternative reading offered for the definition of 'Pp': It is possible that p is true and the sanction is not invoked. I shall suppose that 'the sanction is invoked' is the assigned reading of 'S'.

---

5 It should be noted that Anderson's work was motivated, in part by a desire for a sociological tool to evaluate sociological patterns of norms. See Anderson and Moore [1957].

6 Anderson [1956], p. 171.
The next step in the reduction extends a normal alethic logic so that a normal deontic logic is produced. All this involves is finding suitable axioms to add to a normal alethic logic which are sufficient to produce a normal deontic logic. Anderson does this with a single axiom. It is:

\[(16) \quad M - S\]

and it is read 'It is possibly false that the sanction is invoked.' The reason for making this assumption concerning 'S', Anderson argues, is that if it were not possible for 'S' to be false, then there is nothing one could do including fulfilling all of one's obligations to avoid the invocation of the sanction. And, Anderson says, "The point of choosing penalties in drafting laws...is that the hope of avoiding the penalty might serve as a motivating factor in human behavior; a 'sanction' would serve no such purpose if it were not avoidable."7

The result of adding (16) to any normal alethic logic is a normal deontic logic, and Anderson provides formal proofs to this effect by showing that the desirable and undesirable theorems are or are not present. But the new logic is only an extension of an alethic logic, it is not itself an alethic logic since it contains the unreduced deontic notion of a sanction. At least the notion of a sanction is not an alethic notion.

To complete the reduction we need only replace the new axiom with a theorem found in any normal alethic logic. The search for such a theorem proceeds as follows: Is there an alethic wff that has the

---

7Anderson [1956], p. 171.
same formal characteristics as 'S'? The formal properties of 'S' are that neither 'S' nor its denial are theorems, yet it is a theorem that 'S' is possibly false. The wff which fits this bill of particulars is

(17) \( M \neg p \& p. \)

Thus, (17) is not a theorem, its denial is not a theorem, but

(18) \( M \neg (M \neg p \& p) \)

is a theorem which Anderson derives from a handful of propositional tautologies and the alethic law that 'p \rightarrow M p', a law that must be a theorem of any normal alethic logic. Thus, by replacing 'M-S' with (18) we should have a normal deontic logic which is also a normal alethic logic. But we do not. One further move is necessary.

Anderson replaces both occurrences of the variable 'p' in (18) with the constant 'B'. This is not a trivial move and the difference between 'p' being a variable and 'B' being a constant is all important. If we make this replacement, then 'Pp' is equivalent to

(19) \( M(p \& \neg(M-B \& B)) \)

and 'Op' is equivalent to

(20) \( \neg M(-p \& \neg(M-B \& B)) \)

and it can be shown that (20) logically implies (19) in a normal alethic logic (assuming that 'B' is well-formed). But if the reduction stops at replacing 'M-S' with (18) and if we suppose that 'P' is a unary operator defined in terms of a single variable instead of a variable and a constant as Anderson does, then 'Pp' is defined as

(21) \( M(p \& \neg(M-p \& p)) \)

and since 'Op' is defined as '\neg P \neg p', then to arrive at the alethic re-
duction of 'Op', we must negate (21) and replace every occurrence of 'p' in (21) with '¬p'. The result is

(22) \(-M(\neg p \& -(M\neg p \& p))\).

Notice that the right conjuncts in (21) and (22) differ whereas they are the same in (19) and (20). This difference is the point of replacing 'p' with 'B'.

Now, by DeMorgan's Law, (21) is equivalent to

(23) \(M(p \& (-M\neg p v \neg p))\)

and the equivalence between '¬M-' and 'L' gives

(24) \(M(p \& (Lp v \neg p))\)

which, by distribution, is equivalent to

(25) \(M((p \& Lp) v (p \& \neg p))\)

which, by the fact that 'p v (q \& \neg q)' is equivalent to 'p', reduces to

(26) \(M(p \& Lp)\).

Finally, it is an alethic law that 'p \& Lp' is equivalent to 'Lp', so we arrive at

(27) \(Mlp\)

as a definition of 'Pp'. A similar process applied to (22) reduces 'Op' to

(28) \(Lmp\)

and it should be noted that (28) does not imply (27) in a normal alethic logic. Hence, the constant, 'B', must be employed. Otherwise we have not made a normal alethic logic over into a normal deontic logic. 

To produce a normal deontic logic when 'B' is not added as a constant,
the wff

(29) \[ LHp \rightarrow MLp \]

must be added as an axiom. The point of all of this is that Anderson must either add an axiom like (29) which has "deontic content" or else he must resort to the device of a constant. Either way it seems that some non-alethic device must be added to produce a normal deontic logic, in which case Anderson cannot claim complete success in reducing deontic logic to alethic logic.

4.3 The Systems OM, OM', and OM''

Anderson's names for the three deontic logics he develops suggest that they are deonticized versions of the alethic systems M, M', and M'' which were developed by von Wright. It is now known that they are equivalent to T, S4, and S5, respectively, to which the reader has had a very brief exposure. These deontic logics are generated quite simply. Add 'B' to the primitive syntactical elements of any normal alethic logic, define 'S' as 'M-B & B', 'Pp' as 'M(p & -S)', 'Op' as '-P-p', and 'Fp' as '-Pp' and the deed has been done. Different deontic logics are produced by adding 'B' to the primitive elements of different normal alethic logics.

The only significant difference, as one might expect, between OM, OM', and OM'', is the extent to which they allow the reduction of iterated modalities. So that OM (like T) has an infinite number of deontic modalities, OM' has fourteen ('p', 'Op', 'Pp', 'OPp', 'POp', 'OPOp', 'POPPp', and their negations), and OM'' has six ('p', 'Op', 'Pp', 'OPp', 'POp', 'OPOpp', and their negations).
and their negations).

But whichever deontic logic you choose, it simply is the alethic logic on which it is based plus the constant and suitable definitions. Thus the deontic theorem (required of a normal deontic logic)

\[(30) \quad Op \rightarrow Pp\]

is simply a short-hand way of expressing

\[(31) \quad -M(-p \land -(M-B \land B)) \rightarrow M(p \land -(M-B \land B))\]

which (assuming B is well-formed) is a theorem of any normal deontic logic including T, S4, and S5.

4.4 Some Theorems and a Defense

The reasons in favor of the Andersonian simplification primarily consist of pointing to theorems of OM (as well as OM' and OM'') which seem to capture some intuitively important moral principles and explaining away theorems whose readings at first glance seem paradoxical. Anderson is joined in this task by A. N. Prior who defends Anderson's method of formulating deontic logics in his article "Escapism, the Logical Basis of Ethics."\(^8\)

First, it should be noted that OM rectifies some of the deficiencies of von Wright's first deontic logic. It allows propositional variables as wffs, iterated modalities and mixed modalities. Hence OM can formalize a larger range of uses of 'ought' including relationships

\(^8\)Prior [1958].
between deontic and alethic modalities. Prior and Anderson are both quick to point out the presence of

\[(31) \quad Op + Mp\]

as a theorem of OM which they read as the Kantian principle 'What I ought I can.' Hintikka, however, argues that (31) is too strong. I agree, but I will defer discussion of this issue until Hintikka's work is examined.

Furthermore, Anderson attempts to provide satisfying readings for the paradoxical formulas listed in Chapter I. However, Anderson does not provide a valuation interpretation for his systems which means that there is nothing to which we can appeal to unequivocally decide whether or not his syntactic readings do in fact avoid the paradoxical appearance of the theorems in question. Furthermore, OM does not have among its theorems

\[(32) \quad (0(p \rightarrow q) \& p) \rightarrow 0q\]

which corresponds to the argument mentioned in section 3.6. Thus, according to OM, the argument is invalid. But without a semantics which is related in some clear way to the readings of the formulas of OM, we have no basis for understanding why that argument is invalid or whether it should be.

In sum, there do not seem to be any strong reasons in favor of OM except its larger range of application. And, as we shall now see, there are several strong reasons against it.

4.5 The Use of 'B' Criticized

As Anderson uses 'B', it is an arbitrarily selected propositional constant. Concerning OM, OM', and OM", Anderson says, "Such systems arise regardless of the interpretation placed on 'B'."\(^{10}\) In the next paragraph he says, "It is easy to find an interpretation for 'B' which is consonant with the intended interpretation of 'S' and the deontic constants." That interpretation is, once again, having 'B' express a sanction or penalty. This latter move, since it involves the notion of a sanction, will be discussed in another section. My immediate aim is to discuss the consequences of allowing 'B' to represent an arbitrarily selected propositional constant. Suppose that 'B' stands for 'There is a blue book.' Then the sanction 'S', which is defined as 'M-B & B', could be 'Possibly there is not a blue book but there is.' Moreover, Anderson defined 'F' and 'P' such that 'Fp' is equivalent to

\[(33) \quad L(p \rightarrow S).\]

Hence, on our arbitrarily selected interpretation of 'B',

\[(34) \quad \text{It is forbidden that humans kill one another,}\]

is a short-hand expression for

\[(35) \quad \text{It is necessary that if humans kill one another, then there is a blue book which may have failed to be blue.}\]

But such a consequence seems absurd. Clearly, (34) does not mean the same as, have the same extension as, or follow by any reasonable definitions from (34)\(^{11}\). Hence, if Anderson's systems are to capture ordinary

---

\(^{10}\)Anderson [1956], p. 178.

\(^{11}\)This argument is based upon an example drawn from Castaneda
uses of deontic terms, he must place restrictions on the interpretation of 'B'; it must be interpreted as a sanction or some sort of "bad" consequence. So the next item of business is a discussion of the sanction.

4.6 A Look at the Sanction

It is the use of the sanction in definitions of deontic terms that has drawn the most fire and ire from Anderson's critics. The paper by Castaneda referred to in the previous section contains a series of objections to this move.

Castaneda begins his criticism by correctly noting that Anderson's interpretation of 'S' is quite vague. Castaneda suggests three possible interpretations of 'S':

\( (36) \) A penalty will be inflicted,
\( (37) \) A penalty ought to be inflicted,
\( (38) \) Something bad or wrong will happen.

The first interpretation fails in the face of obvious counter-examples. As Castaneda notes, "It is a commonplace that many criminals die unprosecuted, not to mention those acquitted by a miscarriage of justice."\(^{12}\) Furthermore, (36) runs afoul of a theorem of OM to the effect that the sanction is forbidden:

\( (39) \) \( FS \).

[1960], pp. 41 & 42, which is part of a similar argument directed toward the same conclusion argued for here. His argument, it seems to me, does not clearly make its point.

\(^{12}\)Castaneda [1960], p. 42.
Suppose that the sanction is a prison term. Then (39) reads 'It is forbidden to go to prison' or, perhaps, 'It is forbidden to send someone to prison.' First, it seems implausible, Castaneda says, 'for the established punishments to be forbidden.' It implies, he suggests, that a wrong is committed whenever a law breaker is punished. Consequently, a great deal of wrong doing can be avoided by not punishing. Furthermore, it would seem that a convicted person is obligated to escape punishment. There can be little doubt that interpretation (36) leads to many untenable anomalies.

Interpretation (37) is not a viable alternative for Anderson since the reduction of deontic logic to alethic logic would proceed on the basis of the introduction of an undefined deontic syntactical element. This amounts to no reduction at all.

The third interpretation looks more promising in that it seems reasonable to suppose that an unfulfilled obligation means something wrong has happened or that the world is a little worse off. In other words,

\[(40) \quad Op\]

seems to entail

\[(41) \quad \text{if } -p, \text{ then } (38).\]

But, Castaneda charges, Anderson moves from this reasonable assumption to the position that (41) is itself an entailment. That is for Anderson,

\[(42) \quad -p \text{ strictly implies } (38),\]

\[13\text{Castaneda [1960], p. 43.}\]
in which case

\[(43) \quad -p + (38)\]

is necessarily true. Then Castaneda arms himself with the following principle:

\[(44) \quad \text{A true statement asserting that another statement is logically true must itself be logically true.}\]

and constructs the following reductio against Anderson:

1. If (40) is true, then (42) is true.
2. If (42) is true, then (43) is necessarily true.
3. If (43) is logically true, then (42) is logically true.
4. If (42) is logically true, then (40) is logically true.

Therefore: If (40) is true, (40) is logically true.

But, Castaneda claims, this conclusion is absurd since moral principles must be synthetic

A logical truth is not affected by what we do or by what happens in the world; in the case of deontic assertions, a logically true or analytic deontic assertion is one which cannot be disobeyed, because if prescribes no specific course of action, like 'Either A is permitted or not-A is permitted'.

I agree that the conclusion is unacceptable. A logical truth, after all, is true in all possible worlds. And it is not difficult to think of some logically possible world in which an actual obligation in the real world would not be an obligation or even forbidden. The reader is reminded, however, that I have already argued that even the deontic truth mentioned in the preceding quote is not a logical truth. It is the Principle of Permission discussed at length in the chapter on von

\[14\] Castaneda [1960], p. 45.
Wright. And there I suggested that although it is not logically necessary we may characterize it as being deontically necessary and this distinction suggests that perhaps Castaneda's *reductio* contains an illicit move. Notice the change which occurs between the second premise and the third. Suppose we read 'logically' as 'necessary' in the third premise and the fourth. Then the conclusion should read 'If (40) is true, then (40) is necessarily true' and we may interpret 'necessarily' as 'deontically necessary'. The conclusion is still undesirable. If we pursue the analogy between logical necessity and deontic necessity, then a deontically necessary truth is true in all deontically possible worlds and, if the analogy continues to hold, there are also deontically contingent truths, *i.e.*, obligations and permissions that hold in some deontically possible worlds but not in others. Hence, on this interpretation of deontic expressions (which will be presented more fully in the next chapter), expressions of the form 'Op' are neither logically nor deontically necessary. (Unless 'p' is logically or deontically necessary, then since all deontically possible worlds are logically possible worlds, p is true in all deontically possible worlds. Hence, 'Op' is deontically necessary.)

Looking back at Castaneda's argument, the second premise seems to be a truth of logic while the third and fourth premises are consequences of principle (44). So it seems that the first premise must be rejected. But if the first premise is rejected, then Anderson's definitions of 'O' and 'F' must also be rejected.
But even if we suppose that statements of obligation are necessary and that there is no difficulty in interpreting 'S', it still seems that the "intimate" connection between obligations and sanctions is, in Anderson's hands, too intimate. Surely the anthropologist's inferences mentioned earlier are not grounded in the fact that there is an analytic relation between obligations and sanctions. The relationship between obligations and sanctions is only a contingent relation. It takes no difficulty to imagine our anthropologist running across a society of semi-stoics who believe that they really cannot do anything about evil-doing, yet, unlike Epictetus, they as yet do not have their attitudes under control and so go about their business muttering to themselves about rapes and price-fixing. I suppose that a deontic logician in that society would define a forbidden act as one which has muttering as a necessary consequence, or that there is an intimate connection between obligations and mutterings.

Finally, Castaneda suggests that it is more plausible that the connection, if there is one, between obligations and sanctions goes the other way. Obligations do not seem to require punishment, but punishment seems to require obligations. To institute a system of penalties, there must first be a system of duties. But the reverse does not seem to be true. "A dictator who imprisons and kills at the change of his whim creates no obligations or duties just because of the harm he causes."\(^{15}\)

\(^{15}\) Castaneda [1960], p. 42.
The strongest argument against the use of 'S' in the definitions of 'O' and 'F' was developed by P. H. Nowell-Smith and E. J. Lemmon. Through a reconsideration of a paradox originally discussed and discarded by A. N. Prior they show that the paradox does indeed pose a serious threat to Anderson's contention. This paradox is known as the paradox of the Good Samaritan and arises when a certain reading is given to a theorem of any normal deontic logic in Anderson's sense:

\[(**5)\ L(p \rightarrow q) \iff (L(q \rightarrow S) \rightarrow L(p \rightarrow S))\]

which follows by substitution from the modal law

\[(46)\ L(p \rightarrow q) \iff (L(q \rightarrow r) \rightarrow L(p \rightarrow r)).\]

Prior's reading of (45) is as follows:

"What necessarily implies what necessarily implies the sanction, itself necessarily implies the sanction"; or more briefly, what necessarily implies what is forbidden is itself forbidden. For example, helping someone who has been robbed ("X helps Y who has been robbed" necessarily implies, "Y has been robbed"); but the robbery (being wrong) necessarily implies the sanction; therefore the succor (since it implies the robbery) implies the sanction too, and is also wrong.\footnote{Prior [1958], p. 144.}

Despite the initial implausibility of his reading of (54), Prior believes that the paradox can ultimately be avoided. Nowell-Smith and Lemmon respond that Prior's reading is misleading, if not incorrect, and when corrected is not paradoxical\footnote{Prior [1955], p. 225.}. As usual, the difficulty is in the reading of 'S'. Two possibilities are considered. The first is

\footnote{For the not too terribly interesting details see pp. 293 and 294 of Nowell-Smith and Lemmon [1960].}
'Someone suffers the sanction' and is rejected for the same reasons that Castaneda rejected it. One reason being, among others, that governments seldom have enough resources to capture and punish every violator of every law. The second interpretation is 'Someone ought to suffer the sanction'. This interpretation seems to dissolve the paradox since the original reading suggested that the Good Samaritan ought to suffer the sanction. One cannot object, however, to the consequence that someone ought to suffer the sanction as a consequence of the Good Samaritan's act.

But, alas, the paradox is much worse. Since 'Fp' is defined as 'L(p → S)' we may substitute it in (45) which results in

(46) L(p → q) & Fq) → Fp

which reads very straight-forwardly as 'What necessarily implies what is forbidden is itself forbidden'. Thus, since the Good Samaritan's action necessarily implies the robber's action and since the robber's action is forbidden, the Samaritan's action is forbidden. An adequate deontic logic should not have this result.

Has Anderson given us a reduction? It is interesting to note that the word does not appear in the text of Anderson [1956] but it can be found in the title of Section III, "Reduction of Deontic Logic to Alethic Logic." The answer to the question is that he has not. As noted earlier, Anderson must either add the constant 'B' or an axiom with deontic content. By adding 'B' the resulting system is at least that much removed from the alethic system which underlies the deontic

---

logic. Anderson, however, claims that it does not matter how 'B' is interpreted. Castaneda has effectively argued that it does.
5.1 Deontic Analogs of T, S4, and S5

If the argument of the preceding chapter is correct, then Anderson's series of deontic logics are not acceptable. While their breadth of applicability represents a significant improvement over von Wright's comparatively meager beginning, they apparently go too far; they contain theorems which a deontic logic on that interpretation should not have. A very quick way of generating three deontic logics that are stronger than von Wright's but weaker than Anderson's is by modifying the alethic logics presented in Chapter II in a very simple way. Add to a complete and consistent basis for propositional logic

\[
\begin{align*}
(OA1) & \quad Op \rightarrow Pp \\
(OA2) & \quad O(p \rightarrow q) \rightarrow (Op \rightarrow Oq) \\
(OA3) & \quad O(Op \rightarrow p)
\end{align*}
\]

as axioms and

\[
(R0) \quad \text{If } A \text{ is a theorem, then } OA \text{ is a theorem.}
\]

as a rule of inference. I shall refer to this logic as OT and it is the deontic analog of T. This means that every theorem of T except (A1) and those theorems dependent upon (A1) is also a theorem of OT provided that 'L' and 'M' are replaced with 'O' and 'P', respectively, throughout. Likewise deontic analogs of S4 and S5 may be produced by
adding the deontic analog of the characteristic S4 axiom to OT to produce OS4 and by adding the deontic analog of the characteristic S5 axiom to OS4 to produce OS5.

The semantic models of OT, OS4, and OS5 are just like those for T, S4, and S5, respectively, except that 'L' is replaced by 'O' and 'M' by 'P'. Furthermore, (C.O) has an extra clause not found in its alethic counter-part, (C.L):

(C.O) \[ V(p, W_i) = T \text{ if and only if } V(p, W_k) \text{ for each } W_k \neq W_0 \text{ such that } D(W_k, W_i), \text{ and there is at least one } W_k \text{ such that } D(W_k, W_i). \]

(C.P) \[ V(p, W_i) = T \text{ if and only if } V(p, W_k) = T \text{ for at least one } W_k \text{ such that } D(W_k, W_i). \]

'D(W_k, W_i)' is to be read as 'W_k is a deontic alternative to W_i' and deontic alternative worlds are very much like alethic alternative worlds. Deontic alternative worlds can stand in the same relations as alethic alternative worlds, i.e., (C.Ref), (C.Trans), and (C.Sym) may also apply to deontic alternative worlds. Hence, in addition to satisfying (C.O) and (C.P), OT and OS4 models must satisfy (C.Ref), OS4 models must also satisfy (C.Trans) and OS5 models must satisfy all of these conditions. As before, W_0 designates the actual world.

The semantic readings are changed so that W_i is not a possible world but an ideal world. Whether or not ideal worlds are also possible worlds is taken up in section 6.5. We shall suppose for the moment that an ideal world is a world in which all obligations are satisfied. (C.P) is to be understood as saying that a permitted action is to be
found in at least one world which, according to (C.O) is a world in which all obligations are satisfied, i.e., an ideal world. Hence, on the valuation interpretation of the models of OT, and OS4, and OS5, a permitted action must be consistent with the performance of all duties.

The informal interpretation includes the syntactical reading of 'Op' as 'p is obligatory' and 'Pp' as 'p is permitted' and 'Fp' as 'p is forbidden'. The other elements are read as before. The variables should, in light of previous discussions, be viewed as propositions describing either the description of an action or the state of affairs resulting from the action.

5.2 Iterated Deontic Operators

Since the most notable difference between OT, OS4, and OS5 is the number of non-equivalent modalities to be found in each logic, an obvious question concerns the interpretation of iterated modalities. We shall see that it is not a trivial question. What, if anything, does 'OOp' mean? Does it mean something other than 'Op'? Apply the same questions to 'OPp', 'POp', or even 'POOPOPp'. According to OS4, 'Op' means the same thing as 'OOp', and 'Pp' means the same thing as 'PPp'. Since 'P' is defined in terms of 'O', we shall discuss the 'O' case and assume that our conclusions about 'O' also apply to iterated occurrences of 'P'. There are cases, however, when such reductions as we find in OS4 are not warranted. The common lament "There ought to be a law!" might be interpreted as 'OOp', viz., 'There ought to be a legal obligation that p', where the first 'ought' is not a legal ought. In such
cases, it apparently is felt that the lack of legal sanctions is morally bad. But, if so, then 'OOp' does not seem to reduce to 'Op', since the qualifier 'morally' or 'legally' must carry over, and we cannot carry over the latter since the original lament was over the lack of a law. It is not 'p' that is morally obligatory, it is morally obligatory that there be a law requiring 'p' to be true. Hence the OT theorem 'OOp + Op' fails as a deontic tautology, and OT fails as a deontically adequate logic.

An obvious reply to the preceding reasoning is that 'O' has not been uniformly interpreted. Hence, the argument fails and OT is preserved. Let us consider what happens to the standard systems when 'O' is uniformly interpreted in some full blown sense. Two senses of obligation that seem to be intuitively distinct are moral obligation and legal obligation. Consider, first, legal obligation.

Formulas of the form 'Op' shall be read as 'P is legally obligatory', and the most plausible interpretation of this reading which comes to mind is

(1) There is a law requiring p.

Since 'OOp ≡ Op' is a theorem of OS4 and OS5, in these deontic logics the claim "There ought to be a law" reduces to "There is a law" if we uniformly interpret 'O' according to (1). But this is plainly absurd. Furthermore, the political frailty of legislators occasionally leads them to placate one lobby by passing a law which, among other things, obliges them to pass another law, and yet fail to pass the latter to meet the demands of another more vocal interest group. This often
happens in this country when legislatures pass legislation creating a program or agency with a stipulation that funding must be provided by a certain date, but then fail to pass legislation providing the necessary funds. It is tempting to say that it ought to be the case that if a law is passed requiring a law, then that second law is also passed, which in symbols might be

\[(2) \quad 0(00p + Op).\]

But our uniform reading of '0' rules out (2) as a formalization of that temptation since (2) must begin 'There is a law requiring that...' which is no doubt false. This temptation is accounted for by realizing that what was intended was a moral 'ought'; that the legislators should have complied with the law. At any rate, '00p + Op' is false on a legal interpretation of '0', and since it is a theorem of OS4 and OS5, they must be rejected as acceptable juridical logics.

Our disposal of OS4 and OS5 is a consequence of using (1) as an interpretation of 'Op'. We might suggest an alternative reading of 'Op' which is not as naturalistic as (1) but which still captures the sense of legal obligation:

\[(3) \quad \text{It ought to be that there is a law requiring that } p.\]

This suggestion, however, does not seem to interpret the notion of legal obligation in the same way that our legal system seems to interpret it. And the 'ought' in (4) is probably a moral ought. Iterated moral oughts have their own difficulties. Whereas OT seems, by default, to be an acceptable candidate for formalizing legal obligation, it seems

\[1\text{A term borrowed from Kalinowski [1965].}\]
just as acceptable for moral obligation.

5.3 Iterated Moral Oughts

Consider 'OOp \equiv Op' once again. If I am morally obligated to do 'p', does that imply that the moral obligation itself is morally obligatory? The question could be answered quite easily if we could make sense out of 'OOp'. But comprehension does not come easily when 'It is obligatory that p is obligatory' is rolled around in the mind. There is, however, one clear difference between 'Op' and 'OOp'. The former lends itself quite easily to a rather natural interpretation of what it means for an obligation to be satisfied: If 'p' is true, then the obligation expressed by 'Op' is satisfied. How is 'OOp' satisfied? Is it satisfied when 'Op' is true? Is a moral obligation the kind of thing that can be satisfied by the fact that some other obligation obtains? I do not think so. What follows is an attempt to do two things. First, it is a preliminary examination of the concept of moral obligation, and second, the results of the examination are used to show that 'OOp' is meaningless when 'O' is uniformly interpreted as formalizing moral oughts.

There are two important conditions which must obtain before there is a moral obligation: (1) There must be a moral agent who is bound by the obligation, and (ii) The agent should be able to satisfy the obligation. It is not, after all, such things as rocks, chairs, pencils, and office buildings that are morally responsible for the events in which they play roles. They cannot choose which courses of
action they will follow or what rules they will use to guide their behavior. Rather, the types of entities that can be called to task for their actions are men, angels, and gods. Furthermore, we cannot condemn those who are locked into their destinies by powers beyond their control; the pathological liar, the temporarily insane murderer, the compulsive shop-lifter. No one is morally remiss who fails to save a drowning man because one happens to be miles away and is unaware of the drowning man's difficulties. These practices seem to be quite common and basic to the usual employment of the notion of a moral obligation, and they suggest that (i) and (ii) are conditions upon which the existence of a moral obligation depends.

The first condition may be objected to as follows: Some obligations are universal; they are in effect whether there are any moral agents or not. Even if every last man, woman, and child were wiped from the face of the earth, it would still be obligatory to refrain from killing purely for pleasure. (i) implies that no such obligation is in force. My reply is that

(4) It is obligatory to refrain from killing for pleasure,

is elliptical for something like

(5) (α)(α is a moral agent → (y)(y is an instance of killing for pleasure + α ought to refrain from y))

which, in the strict sense, is not a statement that an obligation does in fact hold. The consequent does that. In the case under consideration (5) is a vacuous universal. But it has the desirable consequence that when moral agents are returned to the world, each and every one of
them will find himself under the obligation to refrain from killing for pleasure.

Another possible counter-example to (i) is suggested by the fact that we sometimes say things like

(6) Policemen ought to be better paid,
or

(7) There ought to be less evil in the world.

These two cannot be disposed of in the same manner as (4) since they do not name a moral agent nor do they seem to suppose that the obligations are binding on every moral agent. Furthermore, how can (7) be binding on anyone? If we suppose that the oughts in (6) and (7) are moral oughts, then we might understand them as follows:

(8) (α)(α is a moral agent ⊃ (y)(y is an instance of raising the pay of policemen ⊃ (it is within the power of α to do y ⊃ α ought to do y))),

(9) (α)(α is a moral agent ⊃ (y)(y is an instance of adding evil to the world ⊃ (it is within the power of α to do y ⊃ α ought to refrain from y))).

(8) and (9) have the distinct advantage over (6) and (7) of giving direction to the practical problem of satisfying the obligations expressed. The main problem is to locate the agent-in-charge, so to speak; someone who has the power to do something about the wages of policemen and evil in the world. In the first case it might be the city council; in the latter, Communists, capitalism and God have all been accused at one time or another.
However, someone who utters (7) might respond to all of this by disclaiming any attempt to specify an obligation on the part of any one, and that their only intention was to simply say something about the world—that it would be better if there were less evil, just as it would be better if prices would drop and our wages would rise. Surely no obligation is intended, and these uses of 'ought' are not typically moral uses of 'ought', i.e., uses of 'ought' which claim that some agent has an obligation to perform some action. But if they are typical uses of 'ought' and if they are not intended to be moral uses, then they are at least evaluative uses of 'ought'. Hence, we can distinguish two senses of 'ought' and consider the possibility of constructing deontic logics to deal with both of them. At the moment, however, I am concerned with the question of whether OT, OS4, and OS5 are acceptable formalizations of moral (as opposed to evaluative) oughts.

The view that conditions (i) and (ii) must obtain in order for a moral obligation to obtain is not entirely eccentric. R. B. Brandt suggests that paradigm contexts of 'ought' have three features in common: "(a) A roughly specifiable service is 'required' of one person. (b) Two parties are involved: the one who is required to perform a service, and the one for whom, or at the bidding of whom, the service is to be performed. (c) A prior transaction, the promise of benefaction, is the source of the relationship." And Brandt muses that "there probably was a time when these three features would have been included in a

2 Brandt [1964], p. 387.
definition of 'obligation'. But to arrive at elements which are now common to all correct uses of 'ought' (there being more uses than the paradigm uses), "we must strike out the notion that obligation is an obligation to someone" and "we must strike out the idea that the bond derives from some prior act or event such as an agreement or the acceptance of a benefaction." Striking out these elements leaves only the elements of a required act and an agent of whom it is required. This is the same as condition (i). Brandt does not mention condition (ii)—but Kant does. Unfortunately, there is sufficient controversy surrounding this condition that I have hedged in the statement of it by using the word 'should' which I felt was vague enough to avoid the problems involved. Despite this disclaimer I believe that 'should' is very appropriate here. My views on this are presented in a more detailed discussion of whether or not ought implies can in section 6.5. For the moment, however, my discussion does not hinge on that later discussion.

The moral of the current discussion is that moral obligations are binding on moral agents who should have the power to satisfy those obligations. If an obligation is satisfied, it is satisfied by the agent exercising that power through the performance of some action. Returning to the formalism, the expression 'Op' expresses an obligation and 'p' in 'Op' expresses the action (or the state of affairs which results from the action) which the moral agent performs to satisfy the obligation, 'Op'. The notion of an act is somewhat loose (is not-going

3Brandt [1964], p. 387.
4Brandt [1964], p. 390.
-to-the-store an act or the failure to act?) and so is the interpretation of 'p', but no matter how loosely we interpret 'act', it does not seem that we can construe an obligation itself as an action. Since 'Op' cannot be conceived of as an action, the description of an action, or the state of affairs resulting from an action (as that phrase is normally meant) which is performable in a way that satisfies '0Op', the interpretation of the first '0' in '0Op' cannot be that of moral obligation. Hence, one cannot uniformly substitute 'morally obligatory' for '0' when one permits the iteration of '0'. So OT cannot be the correct formalization of moral obligation, even though it might work for legal obligation, and OS4 and OS5 will not work for either.

This argument against the standard deontic logics which permit iterated deontic operators might be viewed with mixed feelings. The three logics just rejected are not complicated, their interpretations are initially plausible, and from the point of view of iterated operators, they all contain '0(Op + p)' as a theorem which certainly seems to be a deontic tautology: 'It ought to be the case that obligations are satisfied', even though, as we all know, they are not always in fact satisfied. But it is the notion of satisfying an obligation that did not mix well with iterated operators. Perhaps this problem with '0Op' can be avoided by opting for OS5, since all modalities in OS5 reduce to 'p', 'Pp', 'Op' or their negations, and argue that '0Op' really means nothing more than 'Op'. Or we might tinker with the formation rules to rule out '0Op' but preserve '0(Op + p)'. Unfortunately, there are
other arguments against the adequacy of the standard deontic logics (including P, OH, OM', and OH'') that these last suggestions will not ameliorate.

5.4 The Paradoxes of the Standard Deontic Logics

All of the deontic logics that we have examined have one thing in common. They all contain the standard paradoxes. These are (i) Ross's paradox which turns on the theorem

\[(10) \quad O\ p \land O(p \lor q),\]

(ii) the Good Samaritan paradox which is an attempt to give an unacceptable reading to

\[(11) \quad 0\neg p \lor 0\neg(p \land q),\]

and (iii) the derived obligation paradox which focuses on

\[(12) \quad O\ p \land O(q \lor p).\]

It should be clear that the above theorems are birds of a feather. The second is easily derived from the first by replacing 'p' with '-p' and 'q' with '-q' and applying DeMorgan's Law to the argument of the consequent. The reader should be able to derive (12) from (10). It would seem, then, that they all stand or fall together. Such has not been the case. Fewer writers in deontic logic are willing to accept (12) than (10) or (11). In fact, (12) seemed to be sufficiently damaging to monadic systems of deontic logic to lead von Wright and others to the development of dyadic logics. This move was briefly described in the first chapter. Let us take a closer look at the paradoxes.
5.5 Ross's Paradox

Ross originally used this paradox as a basis for an argument that deontic logic was in principle impossible. His argument is not complex. If it is true that I ought to post a letter (in symbols, 'Op'), it does not follow that I ought to post it or burn it, i.e., 'O(p v q)'. But in virtue of (10) it does follow, and (10) is a theorem of every standard deontic logic. Since these logics contain a principle contrary to clear moral intuitions, these systems are not acceptable.

The opposition to Ross argues that (10) does not lead to conclusions that are as morally degenerate as it first appears. The reason that 'O(p v q)' seems to be an unacceptable conclusion is that we harbor a notion of the satisfaction of obligations that suggests that burning the letter is sufficient to satisfy the obligation described by 'O(p v q)', and it is simply wrong to suppose that burning a letter will satisfy the same obligation that posting it satisfies. However, this explanation of how the paradox arises plays fast and loose with the notions of 'satisfaction' and 'same'. The following three points should be clear: (i) Doing 'p' satisfies 'Op' and 'O(p v q)', but not 'Oq', (ii) Doing 'q' satisfies 'Oq' and 'O(p v q)' but not 'Op', and (iii) 'Op' and 'Oq' and 'O(p v q)' are not the same obligations. Hence, it is possible that posting the letter and burning it might satisfy the same obligation, viz., 'O(p v q)'. But burning the letter does not satisfy the obligation to post it. Lurking behind this para-

---

5Ross [1941].
dox is a feeling that if I satisfy \( \neg(p \lor q) \) then I satisfy any obligations from which it is derived, such as \( \neg p \). But this is clearly not the case if I do \( q \). There is no way that by burning the letter I am able to satisfy the original obligation to post it. Even if it is the case that \( \neg p \) and \( \neg q \) are incompatible and \( p \) and \( q \) are incompatible, which we can easily suppose is the case in Ross's example (in fact, the example may get its apparent force from these assumptions remaining implicit), the paradox does not arise. If I burn the letter, that is, I do \( q \), I have done something that is forbidden which speaks ill of me, not the standard deontic logics, as well as something that prevents me from fulfilling my original obligation to post the letter. It still does not follow that I have fulfilled the obligation to post the letter or even that I am no longer bound by the obligation. In fact, the obligation still obtains.

Now we are at least clear about what the paradox is not. What it is can be illustrated as follows. Ernie promises Bert that he will bake a cake for Bert's birthday and thus incurs the obligation to bake the cake. Then, being a good deontic logician, he realizes that he has a derived obligation to either bake the cake for Bert or paint Bert red. And then it occurs to him that he has yet a third obligation to either bake the cake for Bert or smash Bert in the face with the cake. Well, the big day arrives and Bert shows up for the party. Ernie does three things, he brings out the cake he baked, lets Bert have it in the face, and pours a can of red paint over the already ruffled mixture of Bert
and cake. When Bert finally manages to regain his composure and calmly ask for an explanation, Ernie says, "I was merely doing my duty, Bert. You see, it all started when I promised to bake the cake..." which, given (10) is a perfectly reasonable explanation, but an explanation which neither Bert nor anyone else would find acceptable.

But the paradox gets worse. From 'Op' it follows via (10) that 'O(p v q)' as well as 'O(p v -q)'. Since 'q v -q' is unavoidable, no matter what one does, if one has a single obligation, then any action which is performed satisfies some obligation. But clearly, it would be absurd for Richard Speck, for example, to claim that he was merely doing his duty when he murdered the eight nurses in Chicago. Furthermore, if a person has a single obligation, he has an infinite number of obligations. These consequences indicate that Ross's paradox is an extremely serious obstacle to any consideration of 'p is morally obligatory' as an adequate interpretation of 'Op' for OT, OS4, and OS5.

5.6 A Variation on Ross's Paradox

A weaker version of Ross's paradox is

\[(13) \quad Pp + P(p v q)\]

an instance of which is 'If it is permitted to smoke, then it is permitted to smoke or kill' which seems to justify drawing a rather inflammatory conclusion from a seemingly innocent sign. This paradox lies in the tendency to interpret 'P(p v q)' as 'P(p & q)'. As von Wright points out, 'You are permitted to go to the movie or stay home' implies 'You are permitted to go to the movie and you are permitted to stay
home. But

(14) \[ P(p \lor q) \rightarrow (Pp \land Pq) \]

is not a theorem of any of the standard deontic logics. If it were, we could infer 'Pq' from 'Pp'. Clearly an undesirable consequence. But without (14) 'Pq' cannot be inferred from 'Pp'. (13) seems to be saying that we can. The defect here is not in (13) but in our interpretation of it. Perhaps this point can be clarified just a bit more. According to von Wright, 'You are permitted to go to the movie or stay home' expresses what he calls a 'free choice' permission and it is a rather common use of 'or'. He concludes that there should be deontic logics in which (14) is a theorem and (13) is not.

In a recent work Føllesdal and Hilpinen have argued, correctly, I think, that such a move is unnecessary. A free choice permission can be expressed in monadic deontic logics, they say. The proper formalization of 'You are permitted to go to the movie or to stay home' should be 'Pp \land Pq' rather than 'P(p \lor q)'. To render it as 'P(p \lor q)' is simply a mistake in formalizing ordinary language just as '(x)((Ax \land Ox) \rightarrow Dx)' is an incorrect formalization of 'Apples and oranges are delicious'. Føllesdal and Hilpinen conclude, "There is no need to invent special notations of permission or construct special logics of permission and obligation on the basis of the accidental interchangeability of the words 'or' and 'and' in ordinary language." Hence, even

---

6 von Wright [1968], pp. 21 and 22.
7 Føllesdal and Hilpinen [1971].
8 Føllesdal and Hilpinen [1971], p. 23.
though Ross's paradox makes it impossible to view OT, OS4, and OS5 as explications of moral obligation, this weaker version does not.

5.7 The Good Samaritan Paradox

We have already seen this paradox as it arises in Anderson's deontic logics with disastrous results. The Good Samaritan's action of offering aid to a man who had been robbed entails that a man has been robbed, and robbing is forbidden. We saw in the case of OM that it then followed that the Good Samaritan's action was also forbidden. Clearly unacceptable. The paradoxical character of (11), however, is not as immediately apparent. Making the appropriate substitutions, it reads 'If it ought not be that a man is robbed, then it ought not be that a man is robbed and given aid' which certainly sounds odd. This oddness is the upshot, I think, of the fact that the English sentence suggests that it is forbidden to give aid to the robbed man. If so, then the English is an incorrect reading of (11) since (11) has no such consequence. The strongest comment I have seen in support of the Good Samaritan paradox was made by von Wright who said that this paradox "seems odd, particularly if q is something which we ought to see to that it is the case, when p is the case."9 Well, it may seem odd, but the forest in Figure 5 shows that one can consistently claim '0-p -> 0-(p & q)', '0-p', and '0q'. The forest also indicates that all of the obligations which obtain in the real world can be consistently satisfied only in an ideal world in which 'p' is false. Hence, the Good Samari-

---

9von Wright [1967], p. 137.
FROM

1 \( \sqrt{0 - p + 0 - (p \& q)} \) Assume
2 \( 0 - p \) Assume
3 \( 0 - q \) Assume
4 \( p \) Assume
5 \( \sqrt{0 - p} \) \( \sqrt{0 - (p \& q)} \) Assume
6 \( \sqrt{p} \)

\[ W_0 \]

\[ W_1 \]

1 \( p \) \( W_0 - 5 \)
2 \( q \) \( W_0 - 3 \)
3 \( \neg p \) \( W_0 - 2 \)
X

\[ W_2 \]

1 \( \neg p \) \( W_0 - 2 \)
2 \( q \) \( W_0 - 3 \)
3 \( \sqrt{\neg (p \& q)} \) \( W_0 - 5 \)
4 \( \neg p \) \( \neg q \) 3

FIGURE 5
tan paradox does not pose the difficulties for OT, OS4, and OS5 that it did for OM, OM', and OM". Accordingly, I will not worry about it in the future and will devote my energies to considering Ross's paradox (described above) and the paradox of derived obligation, described next.

5.8 The Derived Obligation Paradox

This paradox has been discussed and worried over more than any of the others. And at one point there was a wholesale abandonment of the standard monadic deontic logics in favor of dyadic deontic logics primarily because no solution to this paradox seemed to be on the horizon. It was expressed in its most deleterious form by its originator, R. M. Chisholm10. Chisholm's very clever argument proceeds in two stages. In the first, he shows that von Wright's $P$ is inadequate to handle contrary-to-duty cases. We may say to a potential thief, "It is forbidden to steal your neighbor's ox," and then add, "But if you do, you should return it and apologize," which in $P$ must be symbolized as

\[(15) \quad 0\neg p\]

and

\[(16) \quad 0(p \rightarrow q).\]

But given

\[(17) \quad 0\neg p \rightarrow 0(p \rightarrow q)\]

which is a theorem of $P$, it is clear that (16) follows from (15), but the same cannot be said for the English sentences of which (15) and (16)

\[10\text{Chisholm [1963].}\]
are the translations. From 'It is forbidden to steal your neighbor's ox,' it does not follow that if you do steal it, that you should return it and apologize. Nor does it follow that

(18) \( O(p \land \neg q) \),

viz., 'If you do, you should not return it and apologize.' These are surely not expected consequences of the eighth commandment. Nor is

(19) \( O(p \lor r) \)

reading \( r \) as 'Steal again and live a life of sin henceforth.' What is needed, Chisholm concludes, is a conditional of the form

(20) \( p \rightarrow Oq \)

to translate 'If you steal your neighbor's ox, you should return it and apologize. Unfortunately (20) is not well-formed in \( P \). Hence, \( P \) is not capable of handling contrary-to-duty conditionals. However, logics in which (20) is well-formed do not fare any better. They are subject to the contrary-to-duty paradox which is the second stage of Chisholm's argument.

Suppose that a certain man ought to go to the aid of his neighbors, which we symbolize as

(21) \( 0p \),

and that it ought to be the case that if he goes, he tells them he is coming,

(22) \( O(p \lor q) \).

But if he does not go, then he ought not tell them he is coming:

(23) \( \neg p \lor 0\neg q \).
But, alas, he does not go to their aid,

(24) \(-p\).

From (21) and (22) we may derive, using the theorem \(\text{0}(p \rightarrow q) \rightarrow (0p + 0q)\),

(25) \(0q\).

From (23) and (24) it follows by Modus Ponens that

(26) \(0\neg q\).

The conjunction of (25) and (26) is a deontic contradiction. Yet the four English statements we have translated seem perfectly consistent.

One might object to the asymmetry of (22) and (23). But if (22) is replaced with

(27) \(p \rightarrow 0q\),

then (21), (27), (23) and (24) is a consistent set (as are the English originals), but it now turns out that (27) is a logical consequence of (24) yet the originals do not intuitively seem to be so related.

Next, if we replace (23) with

(28) \(0(-p + -q)\)

to avoid the asymmetry of (22) and (23), we again have a consistent set of translations, but now (28) is a deontic consequence of (21). I am not rejecting (27) and (28) because they change the moral character of the example. They do not. I am rejecting them because they modify the logical landscape in a way that apparently cannot be reconciled with the example being formalized. These logical relations do not obtain among the originals. There seems to be no nice way out. And since
"most of the situations in which we can assert counter-obligation imperatives are situations in which we can also assert a set of four such statements," standard deontic logics apparently are not adequate to deal with contrary-to-duty imperatives.

Føllesdal and Hilpinen agree and argue that given the interpretation of standard deontic logics it comes as no surprise. According to the ideal world interpretation, in any morally ideal world, the man goes to the assistance of his neighbors and he also tells them that he is coming, i.e., both 'p' and 'q' are true in all morally ideal worlds. Hence, (23) says that in all worlds in which 'p' is false it follows that 'q' is forbidden, so if that world happens to be an ideal world, then 'q' is false. But there are no such ideal worlds; 'p' is true in every ideal world. So (23) does not pertain to ideal worlds, it requires us to look at an "almost" ideal world; a world in which, alas, 'p' is false, but which otherwise is as ideal as it can be. So, to adequately deal with contrary-to-duty imperatives, a second 'ought' operator must be added to standard deontic logics.

Føllesdal and Hilpinen go too far. It is necessary to introduce the notion of an "almost" ideal world to deal with contrary-to-duty imperatives. Their argument depends upon the assumption that Chisholm's asymmetrical translation is correct. I believe it is not.

11 Chisholm [1963], p. 35.
13 The semantic notion of "a world at least as ideal as" is developed in Hanson [1970].
The difference between 'O(p + q)' and 'p + Oq' is not an arbitrary one which can be ignored unless one wishes to avoid various sorts of intuitively implausible inferences as Chisholm has done. In the next chapter we will encounter Hintikka's suggestions for interpreting these two expressions.

We have attempted in this chapter to provide an informal interpretation of OT, OS4, and OS5 as well as a valuation interpretation of their semantic models in terms of ideal worlds. Our informal interpretation of the syntax has attempted to read 'Op' as representing moral obligation. Due to the paradoxes it is quite evident that this combination is not viable. In the next chapter we will look at Hintikka's introduction of quantifiers into deontic logic and we will also look at a new interpretation of the syntax of standard deontic logics. I will argue that this new informal interpretation together with the ideal world semantics introduced in this chapter make standard deontic logics adequate formalizations of the evaluative function of 'ought'.
6.1 Quantifiers

The major contribution of Jaakko Hintikka to deontic logic is an introduction of quantifiers which carried with it a foundation for a clearer understanding of the meanings of deontic expressions. The broad outlines of his suggestions are significant contributions to the advance of deontic logic, but, I believe, the details need correction. The elements of Hintikka's views on deontic logic discussed in the sequel are all found in Hintikka [1971]. This is, in my opinion, the finest paper to be published in this area of inquiry. It is an example of that rare combination of both formal and philosophical sophistication.

In employing quantifiers in deontic logic, there are two reasonable universes of discourse—acts and agents. Hintikka argues that we must quantify over acts. (He does not consider the possibility of quantifying over agents. I will discuss this alternative in the next chapter.) How, Hintikka wonders, are expressions like

(1) \[ a \land b \]

to be interpreted when 'a' and 'b' are intended to be read as the names of actions? There are two possibilities. The first reads (1) as 'On some occasion or other an act with both the property A and the property
B is performed.' On this reading, standard laws of propositional logic fail. First, the rule of conjunction does not hold. From 'a' and 'b' we can no longer infer (1) because 'a' means 'On some occasion or other an act with the property A is performed' and 'b' means 'On some occasion or other an act with the property B is performed' and from these two statements it would be illicit to conclude that an act having both the properties A and B are performed. Quantifiers are clearly of assistance here. A quantified version of the preceding argument would be

\[(2) \quad (\exists x)Ax \]
\[(3x)Bx\]

Hence: \((\exists x)(Ax \& Bx)\)

which we know to be invalid.

Secondly, Disjunctive Syllogism proves to be invalid on this interpretation. According to this interpretation

\[(3) \quad -a \lor b\]

means 'On some occasion or other an act which either fails to have the property A or which has the property B is performed.' From 'a' ('On some occasion or other an act with the property A is performed.') it does not follow that on some occasion or other an act with the property B is performed. In this invalid inference the second premise and the conclusion mention two different acts whereas the first premise, (3), merely mentions one. The act mentioned in the second premise need not be the action mentioned in the first premise. Using quantifiers, the invalidity would again be apparent:
(4) \[ (\exists x)(\neg Ax \lor Bx) \]
\[ (\exists x)Ax \]

Hence: \( (\exists x)Bx \).

The second reading of (1) is 'An act of kind A is done and (perhaps on some other occasion) an act of the kind B is performed.' Under this interpretation, a connective within the scope of a deontic operator does not have the same meaning as the same connective outside the scope of a deontic operator. In other words, to formalize the usual meaning of permission, the negation sign in 'P-a' must be interpreted as meaning something other than what the negation sign in '¬a' means. For example, on this interpretation '¬a' means that it is false that an act of the kind A is done, or, what amounts to the same thing, no act of the kind A (petting one's cat, for example) is done. Translated with quantifiers, it would be \( (\exists x)\neg Ax \). The deontic expression 'P-a' normally means, and must be interpreted as meaning, 'It is permissible that there is an act of not petting one's cat.' However, if the negation sign in 'P-a' is read in the same way it is interpreted in '¬a', 'P-a' would mean that it is permissible that no act is an instance of petting one's cat, and this is much too strong an interpretation for ordinary permission statements. Hintikka claims that the same sort of comment applies to conjunction. He says, "Similarly, 'a & b' means that an act of the kind A is done and (independently of this) an act of the kind B is done, while 'P(a & b)' ordinarily means that one is permitted to do an act with both the properties A and B."¹ I agree that quantifiers

¹Hintikka [1971], p. 66.
will clear up the difficulties of interpreting the negation sign in and out of the scope of deontic operators but I do not agree that the same problem applies to conjunction. Suppose that 'a' means 'These daffodils are moved' and 'b' means 'A hole is dug in the old location of the daffodils.' Then 'P(a & b)' means 'It is permitted to move these daffodils and dig a hole in their old place' which is a common, garden variety of permission statement. Yet it does not mean that a single action having the properties of being an instance of moving daffodils and of being an instance of dumping a load of dirt is permitted. We do not individuate actions as easily as other things, but only a philosopher could think of a case in which a single act would have those two properties. Hence, a uniform interpretation of '&', whether in or out of the scope of a deontic operator requires that we understand it as conjoining two actions. But, it may be objected, a non-quantified deontic calculus which interprets the variables in this way has no way of formalizing

(5) It is permissible to fight and defend one's country, from which it should not be inferred that it is permissible to fight. I suspect, however, that 'and' is being used here idiomatically.

(5) probably means the same thing as

(6) It is permissible to fight in the defense of one's country, and (6) has the advantage of not being an ambiguous conjunction. Depending on the context of 'P(a & b)', it may be translated as either

(7) \[ P(\exists x)(Ax & Bx) \]

or

(8) \[ P((\exists x)Ax & (\exists x)Bx). \]
There is no question that (7) is the correct translation of (6). In general we believe fighting to be wrong, but under certain circumstances we believe it is morally correct—cases of self defense, for example. This is what (6) is saying, that an act of fighting is permissible provided it is also an instance of defending one's country. (6) is not compatible with the suggestion that there be two actions, one of fighting and another act of defending one's country. If quantifiers are not available, then (6) must be translated as 'Pa'.

I have simply taken a way out suggested by Hintikka who points out that a uniform interpretation requires us to "change the meaning of formulae like 'P(a & b)' to say that one is permitted to do A and (independently of this) to do B."\(^2\) But, he says, on this interpretation dire deontic consequences follow: "Under this reading, formulae like 'P(a & b) = (Pa & Pb)' should be provable."\(^3\) But, Hintikka observes, such formulas are not provable in any standard deontic logic. Hintikka's dire consequence is a non-sequitur. The only justifiable reading of the clause "independently of this" on this interpretation is that 'a' and 'b' represent separate actions, not, as Hintikka reads it, that the permissions are independent. They are not. If it is permitted to move these daffodils and dig a hole to repair the pipe, it does not follow that the repairmen are permitted to dig the hole without moving the daffodils, nor does it follow that they are permitted to move the daffo-

\(^2\)Hintikka [1971], p. 66.

\(^3\)Ibid.
dills without digging the hole because daffodils should not be moved unless there is some good reason for doing so. In other words, 'P(∃x)Ax & P(∃x)Bx' does not follow from 'P(∃x)(Ax & Bx)'. Hence, there is no logical absurdity in interpreting the variables as individual actions. By that I mean that we need neither give up clearly valid laws, nor accept clearly invalid ones.

Unfortunately, Hintikka follows his argument to the bitter end. Persuaded by it, he drops the interpretation of (1) as a conjunction of the descriptions of the two individual actions and returns to the first interpretation, that a single act having both properties is done. Now, in order to preserve Conjunction as a rule of inference, he must introduce the semantic device of a situation. He says,

> We have to take formulae like 'a & b' to be concerned with some (arbitrarily chosen) particular situation. In other words, 'a & b' should be read: 'An act of the sort A is done in the particular situation we are considering, and an act of the sort B is done in the same situation.' Under this interpretation alone is one entitled to infer from 'An act of the sort A is done' and 'An act of the sort B is done' that an act of both the sort A and the sort B is done.'

It seems to me that the "particular situation" would need to be narrowly circumscribed for this conclusion to follow, and I do not see that the choice of how far to extend the situation can be made arbitrarily. In order for the conclusion to be inescapable (which it must be) the "particular situation" would have to be limited to a single act. It might be urged, though, that Hintikka means to limit the "situation" to a particular time and place. Is it possible for more than one action

---

4Hintikka [1971], p. 66.
to be performed at a particular time and place? If not, then Hintikka has his wish and the inference he is concerned with is valid. If it is possible for there to be more than one action in a single place and time, and it certainly seems that this is possible (e.g., eating a sandwich, looking at a picture, and tapping a foot all at the same time), then the inference is invalid. It is valid only if one supposes that there is something about a situation that limits it to containing only a single action. However, I believe that I have shown that no such restrictive device is necessary by showing that (1) may be interpreted as conjoining the names (or descriptions) of two individual actions and that this does not have the undesirable consequences which Hintikka believes it does.

Even so, the calculus remains inferior to what we could have if quantifiers were introduced. The difficulties with negation would be avoided and we could express formally claims that a single act individual has several properties. This does not mean that an adequate deontic logic requires quantifiers. But it does mean that a deontic logic which applies deontic operators to act individuals can capture more logical distinctions with quantifiers.

6.2 The Interpretation of HDL

Hintikka's directions for representing statements of obligation, permission, and prohibition with wffs of HDL (Hintikka's Deontic Logic) are cursory at best. In this section I shall show that the few translations he provides cannot all be correct. Then, building on Hintikka's
errors, I will show how the wffs of HDL must be interpreted given the view argued for in the previous section that the variables must represent statements whose subjects are act individuals. Hence, in HDL, a predicate function will be interpreted as saying that a certain act individual has a certain property. But first, let us look at Hintikka's proposals.

Hintikka believes that a statement like 'Apologizing is forbidden,' should be expressed as

$$(9) \quad (x) F A x$$

which is to be read as 'Every action ought not be an instance of apologizing.' In a non-quantified deontic logic, 'Fp' is equivalent to 'O-p' and statements of prohibition and obligation are translated as 'Fp' and '0p', respectively. Notice that 'F' has been replaced by '0-' and 'p' has been replaced by '-p' and the resulting double negative has been eliminated. So, if we follow this pattern, replacing 'F' with '0-' and '-Ax' with 'Ax' gives (ignoring the quantifier)

$$(10) \quad (x) O A x$$

as the translation of 'You ought to apologize.' But (10) is obviously too strong. It says that every action ought to be an instance of apologizing. But, ordinarily, 'You ought to apologize' suggests merely that there ought to be at least one action which is an instance of apologizing. Instead, it seems that there ought to be at least one action which is an instance of apologizing. Hence,

$$(11) \quad O(\exists x) A x$$

is the correct formalization of 'You ought to apologize.' Hintikka does
not say why \((\exists x)\Box Ax\) is rejected.

To produce the formal representation of 'Apologizing is permitted', Hintikka again claims that the non-quantified relationships between deontic notions cannot be uncritically carried over to a quantified deontic calculus. If we suppose that permission is related to prohibition "in the way it is usually assumed to be related," viz.,

\[(12) \quad Pp \equiv -Fp,\]

then

\[(13) \quad (\exists x)PAx\]

must be the formal explication of permission. And an appeal to the "usual" relationship between permission and obligation, viz.,

\[(14) \quad Pp \equiv -0-p\]

produces

\[(15) \quad P(x)Ax\]

as the proper form of permission. Hintikka rejects them both and on the basis of a brief argument, chooses, instead,

\[(16) \quad (x)PAx.\]

The argument will be examined presently. Unfortunately, Hintikka's explanation of these "usual relations" and how (13) and (15) are derived from (12) and (14) is almost as tersé as my paraphrase of it. Unfortunate, because those usual relations found in non-quantified deontic logics have a strong intuitive foundation, and if one is going to reject them, some argument is necessary. I shall argue that Hintikka's formal representations of permission and prohibition are incorrect. I believe

\[\text{Hintikka [1971], p. 64.}\]
this to be the result of a lack of appreciation on Hintikka's part of the difference between 
'(x)OAx' and 'O(x)Ax'.

The defect to be found in Hintikka's translations is that they do not properly represent the logical relations which exist between obligation, permission, and prohibition. He offers few reasons for his choices but there are good reasons against them. To see what these relations are, a decision procedure is needed. Rather than present Hintikka's rules and semantics, I will simply say that they amount to the rules for OS4 presented in the previous chapter, plus rules for quantifiers. In any given world, whether it be the original world or an ideal world, the usual tree rules for quantifiers apply with only one slight modification necessary to coincide with the following conditions:

(C.E) If $V((\exists x)\phi x,W_i) = T$, then $V(\phi_\mu,W_i) = T$ for some member $\mu$ in $W_i$. Otherwise $V((\exists x)\phi x,W_i) = F$. 
(C.U) If $V((x)x,W_i) = T$, then $V(\phi_\mu,W_i) = T$ for all members $\mu$ of $W_i$, otherwise $V((x)x,W_i) = F$.

In constructing models which fulfill these conditions, Hintikka believes there is a precaution which must be observed: Unless forced to import an individual action from one world to another by rule (C.P), there is no reason to suppose that an individual which exists in one world also exists in a world which is a deontic alternative to it. In fact, Hintikka believes there are reasons to think that in many cases they do not. I shall illustrate the restriction, share Hintikka's argument with
you, and then comment favorably on it.

In the forest depicted in Figure 6, the individual action 'a' is moved from \( W_0 \) to \( W_1 \). This is in accordance with (C.P) and Hintikka finds it perfectly unobjectionable. But the situation is different in the next forest of Figure 7. Hintikka believes that line 2 of \( W_1 \) in Figure 7 is an illicit application of (C.O). Hence,

\[
(17) \quad (\exists x)OAx \rightarrow O(\exists x)Ax
\]

should not be a theorem. Hintikka's argument is this: "If there is, under the actual course of events, an act that ought to be an instance of forgiving a trespass, it clearly does not follow that there ought to be, under any deontically perfect course of events, an act of forgiving --and, hence, presumably, also an earlier act of trespassing."^6

Furthermore, if (17) is deontically valid, then so is

\[
(18) \quad (\exists x)O(\exists y)Pyx \rightarrow O(\exists x)(\exists y)Pyx
\]

which says, for example, that "from the fact that there is an act (which therefore 'exists' in the usual non-temporal sense of existence) which ought to be punished, it does not follow that there ought to exist a punishable act."^7 I accept Hintikka's argument, but several questions need to be considered: (i) Can the discussion of quantifiers in deontic logic profit from what is known about quantified alethic modal logics? (ii) Should we be uneasy about the fact that the rules apparently differ in treating 'PAa' and '0Aa' (re-examine the preceding two forests in Figure 6 and Figure 7)? (iii) Can anything more be said about what

---

^6 Hintikka [1971], p. 102.
^7 Ibid.
\begin{align*}
\text{FIGURE 6} &\quad \text{FROM} 124 \\
W_0 &\quad \begin{array}{|l|}
\hline
1 & \sqrt{-\left(0(x)Ax + (x)OAx\right)} \text{ Assume} \\
2 & \sqrt{0(x)Ax} * \hspace{2cm} 1 \\
3 & \sqrt{-(x)OAx} \hspace{2cm} 1 \\
4 & \sqrt{(3x)P-Ax} \hspace{2cm} 3 \\
5 & \sqrt{P-Aa} \hspace{2cm} 4 \\
\hline
\end{array} \\
W_1 &\quad \downarrow \\
\text{FIGURE 7} &\quad \text{FROM} \\
W_0 &\quad \begin{array}{|l|}
\hline
1 & -Aa \hspace{2cm} W_0-5 \\
2 & (x)Ax \hspace{2cm} W_0-2 \\
3 & Aa \hspace{2cm} 2 \\
\hline
\end{array}
\end{align*}
It means for an act to exist in any world? (iv) Does it make sense to speak of the same act existing in two different worlds? I will discuss (i) and (ii) now and take up the others after assessing Hintikka's proposed translations.

The combination of alethic operators and quantifiers is a philosophically volatile mixture. Much of the controversy centers on what is known as the Barcan Formula, an instance of which is

\[(19) \quad (x)\Box A_x \rightarrow \Box (x)A_x.\]

The Barcan Formula is falsifiable if things may have different properties from world to world or if there are things in possible worlds which do not exist in the actual world. The Barcan Formula is not germane here, but it does have a deontic analog:

\[(20) \quad (x)\circ A_x \rightarrow \circ (x)A_x\]

which Hintikka must either reject or else lift his restriction on importing act individuals into ideal worlds. Given that Hintikka specifically rejects (20)\(^9\) it follows that an ideal world may contain acts which do not exist in any of its ideal worlds. Given this it can be shown that either

\[(21) \quad (x)A_x \rightarrow A_a\]

is invalid or that not every atomic formula is either true or false in every world\(^10\). In HDL, (21) is valid. Hence, the semantics of HDL

---

\(^8\) See Hughes and Cresswell [1968], Chapter 10 for a discussion of the formal aspects of this issue and references to discussions of the philosophical aspects.

\(^9\) Hintikka [1971], p. 103.

\(^10\) Hughes and Cresswell [1968], pp. 178 and 179.
must not specify a truth value for every atomic formula in every world. This seems intuitively acceptable. Suppose that 'Aa' represents the attempted assassination of George Wallace. In the actual world 'Aa' is true, but need it have a truth value in an ideal world? Must 'Aa' be false in an ideal world because 'a' does not exist in an ideal world? In this case, it is tempting to say so, but if 'Aa' represented a morally neutral action, the temptation wanes. For example, suppose that 'Aa' says that 'a' is an instance of picking apples. The truth value of 'Aa' in an ideal world need not be determined one way or the other. We must let the deontic statements of the actual world be our guide of what is true and false in its deontic alternatives. So, if 'O-Aa' is true in the actual world, then there is at least one ideal world (the one in which 'Aa' is permitted) in which 'a' exists, and in that world 'a' is an instance of picking apples. But if 'Aa' is neither permitted, nor obligatory, nor forbidden, then there is no telling whether 'a' is a member of an ideal world or not, and the truth value of 'Aa' in all ideal worlds is undetermined. In other words, if (OAAa,Wj) = T, then (Aa,Wk) = t for any Wk such that D(Wk,Wj) and 'a' is a member of Wk. So much for the first question posed above.

In answer to the second, examine the forest of Figure 8. Line 1 of W1 is obtained by (C.O) which says that there is at least one ideal world in which what is obligatory is true, and W2 is obtained by (C.P). This forest has two open paths. But the forest of Figure 9 closes. In Figure 9, line 2 of W1 is gotten by (C.O) since the act 'a' does
FROM

1  OAa \neq \text{ Assume}
2  \sqrt{P(x)}-Ax  \text{ Assume}

W_1 \downarrow \hspace{2cm} W_2 \downarrow

1  Aa  W_0-1 \hspace{2cm} 1  (x)-Ax  W_0-2

FIGURE 8

FROM

1  OAa \neq \text{ Assume}
2  (x)P-Ax  \text{ Assume}
3  \sqrt{P-Aa}  2

W_0 \downarrow

1  -Aa  W_0-3
2  Aa  \text{ W}_0-1
X

FIGURE 9
exist in $W_1$ by (C.P). Hence, (C.O) is a broader rule than (C.P) with Hintikka's restriction even though the examples used to present the restriction suggest that (C.O) is a weaker rule. 'OAa' gives 'Aa' in all worlds of which 'a' is a member, while 'PAa' gives 'Aa' only in at least one world which has 'a' as a member. These two examples raise the question of how 'P(x)Ax' and '(x)PAx' are to be interpreted. Clearly, as these models show, there must be a difference. But before we consider what that difference might be, after this lengthy aside, we should return to the matter at hand; an assessment of Hintikka's translations of obligation, premission, and prohibition. Then I shall discuss the interpretation of expressions of other sorts in HDL.

We are now in a position to determine the logical relations between the important modalities of HDL. To see what these relations are I offer the Square of Implication for Quantified Deontic Modalities (see Table 5). The arrows indicate the direction of implication. I have called it a "Square" of implication because it closely resembles the Square of Opposition for these same modalities (see Table 6).

Both squares were developed the hard way: various combinations of modalities were tested via forests and the results recorded in a list of deontically valid expressions as they were tested. Finally, a pattern began to emerge and the squares were developed since they display the logical features of HDL much more perspicuously than a list of valid formulas. Each relationship exhibited by the squares has been subjected to the forest test. In addition to the relationships labelled in the Square of Opposition, diagonally opposed expressions on the outer square are also contradictories.
TABLE 5

SQUARE OF IMPLICATION FOR QUANTIFIED DEONTIC MODALITIES

TABLE 6

SQUARE OF OPPOSITION FOR QUANTIFIED DEONTIC MODALITIES
The immediately striking thing about the Square of Opposition is that Hintikka's translations of obligation ('0(∃x)Ax') and prohibition ('(x)0Ax') are not formal contraries whereas 'A is forbidden' and 'A is obligatory' may surely not both be true. Hintikka's translations are obviously in error. To express 'A is obligatory' and 'A is forbidden' we must choose either '0(∃x)Ax' and '0(x)0Ax' or else '(∃x)0Ax' and '(x)0Ax'. Which pair should we choose? What is the difference between them?

Initially, there does not seem to be any clear cut answer. Conceptually, we do not deal with actions nearly as easily as we do 'things' in general. It is much more difficult to individuate actions; it is not easy to tell where one ends and another begins; actions do not last as long as most other things, hence, it is more difficult for them to be the objects of cognitive attitudes; once done, they are not susceptible to change and manipulation; relations between them seem to be limited to temporal, intentional, and moral relations. What is it to say that an action exists? Do actions performed centuries ago exist? Do actions yet to be performed exist? I will attempt to answer these questions within the framework of the assumptions already made about the interpretation of HDL, that its universe of discourse will always be act individuals, and that one act need not exist in one world even if it exists in another, and that all expressions of HDL are atemporal in the usual way. I believe that these assumptions are perfectly plausible and I will not attempt to fill in the details.
When applying a quantified calculus, we may always restrict
the range of discourse in any way that suits our purposes as long as
the same logical relations obtain. Are we free to make the same stipu-
lations when quantifying over actions? We may specify the range of
discourse over actions in any of several ways. First, we might try all
actions, past, present, and future, or, all actions performed on July
4th, 1776, or all of the actions of Socrates from birth to death, or
all of the actions of a living person from birth to death including the
actions which that person has yet to perform, or all past actions or
all future actions. All of these suggestions are reasonable as long as
the rules of inference hold and the formulas are readily interpreted.
I believe that both of these conditions hold for all of the domains
listed above, but, I find a few of these suggestions not entirely natur-
al, viz., those which include future actions. I am not going to argue
that ranges of discourse which include future actions should not be em-
ployed. But unless there is a clear special purpose for doing so, the
ordinary meanings of deontic statements are best captured by not doing
so. This unnaturalness is the result of not being clear about what it
means for an action to exist. I hope to remedy this.

Let us suppose that future actions exist and let us suppose
that at the end of the world when humanity has breathed its last that
it is false that there is an instance of Al repaying his debt. This is
all that can be meant by

\[(22) \quad -(\exists x)Ax\]

no matter when (22) is uttered. It says that there never has been nor
will there ever be an Instance of Al repaying his debt. I have left out the phrase "nor is there now" for the sake of economy. I am supposing that all actions are either past or future. Any putative counterexample, an action which is neither future nor past but happening now, is hereby declared incomplete and until it is completed it is a future action—an action whose character could be influenced by moral imperatives or advice. On the supposition that future actions exist, how are we to interpret

(23) \(0(\exists x)Ax\)?

It can be nothing more than a lament; a wish that the universe had been constituted differently. Specifically, we can not interpret (23) as advice to Al since on our present interpretation of what it is for an act to exist, (22) says that Al is destined to never repay his debt. A much more natural assumption is that the stock of all existing actions consists of past actions but not future actions. If one wishes to add another action to the supply of existing actions, then one just does it—the action and adding to the supply. Hence, (23) positively values adding an instance of A to the present set of actions, and a statement like 'You, Al, ought to repay that debt!' is advocating that Al add an instance of A to the set of all actions. On this interpretation of what is and is not an existing act, even if (22) is true, (23) can be interpreted as an entreaty directed to Al urging him to do a certain sort of act. If he does, then (22) is falsified. It was, in this case, an "accidental" universal. On this interpretation

(24) \((\exists x)0Ax\)
is an after the fact evaluation; a claim that an action which has been
done ought to be an instance of A. The action may or may not be an
instance of A, but it does exist. (23) makes no such existential
commitment.

I conclude, then, that the typically moral 'ought' which posi-
tively evaluates a certain sort of action being done cannot be formaliz-
ed as (24) since (24) is essentially retrospective. Claims about what
ought to be done must be translated as (23). But if we suppose that an
'ought not', a prohibition, must be at least the contrary of an 'ought',
then Hintikka's suggestion to translate prohibition as '(x)\neg A x' must
be rejected because its contrary, we have decided, and Hintikka agrees,
is an inappropriate vehicle for formalizing oughts.

Must Hintikka's choice of form for permission be revised as
well? I think so. My quarrel with

\[(25) \quad (x)PAx\]
is with both the quantifier and its position, but our difference over
its position is the most crucial difference. Most, if not all, state-
ments of permission have restrictions of one sort or another built into
them. When children seek permission to do various things, the subse-
quent permission statement is not always, or even often, interpreted as
a universal but rather that on this occasion it is permissible, for
example, to stay up past nine p.m. 'All right, but just this one time,'
cannot be understood as saying (25) nor as (3x)PAx since the statement
is about an action yet to occur. The correct form is the only one left,

\[(26) \quad P(3x)A x.\]
There are times, however, when a universal quantifier seems required, as, for example, in 'Consuming alcohol is never permitted.' It is because of this case that Hintikka chose (25) over

\[ (27) \quad P(x) \forall x \]

as the standard formalization of permission. "In most countries, it is legally permissible to consume alcoholic beverages; but it is absurd to say that this permission is logically the same as a permission to drink on every occasion."\(^{11}\) Hintikka apparently views such behavior as morally wrong. Even so, the case he presents concerns legal permissibility and I assume that it is legally permissible to do nothing but drink alcohol in those countries where it is legally permitted provided that one can afford it and does not break any other laws (public drunkenness, etc.). Furthermore, I do not see that this passage indicates that the distinction between (25) and (27) is what Hintikka thinks it is. If I am correct about the manner in which quantifiers over actions should be interpreted, then (25) says that every action is an action which may be an instance of consuming alcohol, and, incidentally, there is no guarantee that some falsifying case will not arise tomorrow, especially if the WCTU comes back into power. If one reads the quantifier as 'Pick any action and it will be such that...' then there is nothing preventing an alcoholic saying "Good, I pick them all," which is the sort of thing that Hintikka would apparently like to have his deontic logic rule out. I do not believe that it does. Nor do I believe that it should.

Consider a sign which says "Smoking is permitted." It would be incorrect to formalize the message of the sign as (27) or (25) if 'Ax'

\(^{11}\)Hintikka [1971], p. 64.
is interpreted as 'x is an instance of smoking' for the reason that such signs are often located in areas proximate to others where one generally not smoke. Instead, if we are to fully appreciate the point of the sign, 'Ax' must be read as 'x is an instance of smoking in this area.' It seems to me that the message of the sign is then quite adequately captured by (27), i.e., one may sit in this area and do nothing but smoke. If there are restrictions attached to a permission, then we must specify them by building them into the predicate or using a conditional within the scope of the quantifier or using an existential quantifier. When one is not sure what the restrictions are, or if one is not sure whether the permission is universal, then the safe bet is to formalize the permission as (26) since the Square of Implication for Quantified Deontic Modalities tells us that (26) is the weakest of the modalities.

If we adopt (26) as the general form of permissions, then the relations between 'A is obligatory' and 'A is permitted' and 'A is forbidden', are exactly those between 'O', 'P', and 'F'. Hence, contrary to what Hintikka believes, the relations between obligation, permission, and prohibition are not more complex than anyone had previously thought.

What could have motivated Hintikka to adopt the views about his deontic logic that have just been criticized? Look again at the last quotation from Hintikka quoted above. He uses the phrase 'on every occasion' which implies that he takes the values of the variables to be not actions but occasions. I believe that the inclusion of the phrase
"on every occasion" is motivated by the argument we examined earlier regarding the interpretation of 'S' and the necessity of quantifiers. There, he concluded that the act variables in a conjunction must be related to some "particular situation". This was done to avoid undesirable theorems containing conjunctions within the scope of 'P'. Hence, it is only natural to construe the universal quantifier as a claim about all occasions. Furthermore, it is quite natural to read a universal quantifier to the left of 'P' as 'any' and to read it as 'every' when it falls to the right of 'P'. I have rejected Hintikka's conclusion that in addition to speaking of individual actions we must also consider "particular situations", and if quantification over act individuals is taken seriously, we are forced into interpreting HDL in the manner which I have been urging throughout this section. Hence, his original intention, to quantify over actions, is, I believe, a most fruitful one. And given the semantics of HDL, the domains of all worlds, actual and ideal, are act individuals. The values of the variables of quantifiers to the left of deontic operators (e.g., '(x)OAx') are act individuals which exist in the actual world (completed or past actions). The values of the quantifiers which appear to the right of deontic operators (e.g., 'O(x)Ax') are act individuals which exist in (one or more) ideal worlds which, in the case of obligations, it is hoped will be realized in the future in the actual world. It is a mistake to suppose that the domain of quantifiers to the right of deontic operators consists of future actions. Future actions are not existing actions in the actual world or in ideal worlds. This mode of interpre-
ting combinations of deontic operators and quantifiers will be utilized in the last chapter.

6.3 Deontic and Logical Consequences

Despite Hintikka's unfortunate interpretations of quantifiers, he points to a classical distinction in ethics which he believes is formalizable in HDL. He uses this distinction to criticize John Searle's well known paper "How to Derive Ought From Is."\(^{12}\) This distinction is the oft cited difference between absolute obligations and prima facie obligations. It is a species of the difference between deontic consequences and logical consequences. First, I shall explain this latter distinction.

The distinction is simply this: If \(p \land q\) is valid (i.e., a tautology), then \(p\) logically implies \(q\). If \(\roll(q \to p)\) is valid (i.e., a deontic tautology), then \(p\) deontically implies \(q\). The difference is between whether one is discussing any arbitrarily chosen possible world or only deontically perfect (ideal) worlds. The two notions have been confused. A. N. Prior thought that

\[
(28) \quad (0p \land (p \to 0q)) \to 0q
\]

should be a deontic theorem\(^{13}\). Using the terminology introduced above, Prior is claiming that (28) is a logical consequence. It is not. Its plausibility stems from the fact that it is a deontic consequence. In

---

\(^{12}\)Searle [1966].

\(^{13}\)Prior [1955], p. 225.
other words,

\[(29) \quad 0((p \& (p \rightarrow q)) \rightarrow q)\]

is a deontic theorem whereas (28) is not. The following forest explains how (28) can be false:

\[W_0\]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>(\neg((0p &amp; (p \rightarrow q)) \rightarrow q)) Assume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>(0p &amp; (p \rightarrow q)) 1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>(\neg q) 3</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>(p \neg q) 2</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>(0p \neg q) 5</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>(p \rightarrow q) 2</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>(-p) 0q</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>(X)</td>
</tr>
</tbody>
</table>

\[W_1\]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>(-q) (W_0-4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>(p) (W_0-5)</td>
</tr>
</tbody>
</table>

It is false when the obligation to do 'p' is left unfulfilled. Then, of course, the obligation to do 'q' does not follow (since 'p' is not true).

The most fruitful consequence of this distinction is in the difference between

\[(30) \quad p \rightarrow q\]

and
which, Hintikka claims, captures the difference between actual obligations and prima facie obligations, respectively. Traditionally, the distinction views prima facie obligations as obligations that may be overridden by higher obligations. For example, I create an obligation to be in New York City by promising you that I will meet you there. But in the meantime, a close relative has fallen ill obligating me (more strongly, let us suppose) to be elsewhere, an obligation which takes precedence over the prima facie obligation to meet you. An actual obligation cannot be overridden in this way. A prima facie obligation, at times, appears to be no obligation at all since it does not help to tell me what I actually ought to do. But, still, things are not completely all right if I must break the promise to meet you. Hintikka feels that he can account for all of this. If 'p' represents 'I promise to meet you in New York' and 'q' stands for 'I meet you in New York', then if I promise to meet you, from (30) it follows that I ought to meet you, hence, an actual obligation. But from (31) nothing follows except that in an ideal world I do meet you. But this world is not morally ideal. Duties often conflict and my actual duty might ultimately lie somewhere else. And the vague feeling that something has gone wrong when a prima facie obligation has been violated is accounted for by noting that "something takes place that would not happen in a deontically perfect world."14

---

14Hintikka [1971], p. 93.
Hintikka applies the distinction between deontic consequences and logical consequences to a well known argument of John Searle. Put quite simply, Searle argues that it is possible to infer that one has an obligation to keep a promise from the factual premise that one has made a promise. The obviously needed intermediary step is the "tautological" principle that promises ought to be kept. Let us have 'p' represent the statement that a promise is made and 'q' stand for the claim that the promise is kept. Searle's argument has this form:

\[ \text{(32)} \quad p \]

Promises ought to be kept.

Hence: Oq.

The second premise may be represented by either (30) or (31). Hence, the complete formal representation of (32) is either

\[ \text{(33)} \quad p \]

\[ p \rightarrow Oq \]

Hence: Oq,

or

\[ \text{(34)} \quad p \]

\[ O(p \rightarrow q) \]

Hence: Oq.

Note that (34) is invalid, while (33), being an instance of modus ponens, is valid. Hence, if Searle's argument is correct, we must suppose that the obligation to keep one's promises is an absolute obligation. But, clearly, our obligation to keep promises can be overridden by other com-

\[ ^{15}\text{Searle [1964].} \]
peting obligations. Thus (33) fails as badly as (34) unless one can strengthen 'p' in such a way that the second premise is analytic. To do this Searle uses a ceteris paribus clause conjoined to 'p'. The argument now looks like this:

\[(35) \quad p \land cp\]
\[(p \land cp) \rightarrow Oq\]

Hence: \(Oq\)

reading 'cp' as the ceteris paribus clause. Notice that 'cp' must be conjoined to 'p' in the first premise in order to preserve the validity of the argument. In this case we may easily suppose that there are no counter-examples to the second premise and that it is analytically true. But the first premise, Hintikka claims, due to the 'cp' clause, is now a normative proposition. Hence, this is still not a case of deriving an ought from an is. It is quite true that there are cases in which there are no competing obligations (the content of the 'cp' clause) but to claim that this is so is a moral claim. Saying that something is neither obligatory nor forbidden is just as much a normative claim as saying that it is either obligatory or forbidden.

Searle anticipates this last criticism and responds that the 'cp' condition may be made part of the conclusion. Hintikka responds that Searle's new argument will be either

\[(36) \quad p\]
\[p \rightarrow O(cp \rightarrow q)\]

Hence: \(cp \rightarrow Oq\).

or
(37) \[ p \]
\[ p \to (cp \to 0q) \]

Hence: \( cp \to 0q \).

In comment, the second premise of (36) is false. It cannot follow from a factual claim about this world that something is true in all ideal worlds unless '0(cp + q)' is analytic. If it is analytic, then the conclusion follows independently of 'p', and Searle gains nothing. The conclusion of (37) prevents this form of his suggestion from doing what he hopes to do, too. The conclusion says that if one normative statement is true (the 'cp' clause), then another normative statement is true, and this is far from claiming that an actual obligation exists; that an ought follows from an is. Either way, Searle still fails to bridge the fact-norm gap.

It does seem though, that Searle ought to be right. And, indeed, he ought. What this means is that he is on to something which is not a logical consequence, as he claims, but a deontic consequence.

The premise that promises ought to be kept was seen to be false when interpreted as '(p + 0q)'. But if we view it as a deontic consequence, i.e., as '0(p + q)', then it certainly seems analytic that in all ideal worlds promises are kept. We noted, however, that (34) is an invalid argument; that it is not a logical consequence. But it does represent a deontic consequence. In other words,

(38) \[ (p \& 0(p + q)) \to 0q \]

is not a valid formula, but

(39) \[ 0(p \& 0(p + q)) \to 0q \]
is a valid formula. Hence, if Searle's claim is interpreted as saying that an actual obligation can be derived from an 'is', he is wrong. If he is claiming that a prima facie obligation can be derived from an is, then, of course, he is right.

6.4 A Critique of HDL

I have argued in section 6.1 that Hintikka's arguments in favor of quantifiers can be avoided. I believe that it has been successfully shown that one may interpret non-quantified variables as names of acts and avoid the difficulties that Hintikka incorrectly raises. But this is not to say that a non-quantified version of HDL is without problems. HDL without quantifiers is deductively equivalent to OS4, and in the previous chapter it was argued that if we suppose that OS4 is about duties and obligations, then it is an unacceptable deontic logic in virtue of Ross's paradox, the paradox of derived obligation, and the contrary-to-duty paradox. If Hintikka interprets HDL in the same way, then HDL is unacceptable for the same reasons (except that I will show in section 6.6 that HDL avoids the contrary-to-duty paradox).

Hintikka does interpret HDL as being about duties and obligations. He says,

A deontic alternative to a given world, say the one described by μ, is intended to be a kind of deontically perfect world, when viewed upon from the point of view of the norms specified by μ. This requirement of deontic perfection means that all the relevant duties are fulfilled in the alternative world. They naturally include all duties obtaining (in virtue of the normative system specified by μ) in the alternative itself,
no matter whether they obtain in the world described by μ or not.\textsuperscript{16}

Hence, due to the paradoxes enumerated above, HDL is not an acceptable deontic logic. The difficulties of HDL, however, do not end here. In a recent paper\textsuperscript{17}, Richard Purtill contends that "it is very difficult to say whether Hintikka's machinery of 'deontic alternatives' is viable, for too many unanswered questions arise with regard to that machinery."\textsuperscript{18} First, ideal worlds have the undesirable feature that if there are no people in an ideal world, then there are no obligations. If there are no obligations (i.e., \(-\text{Op}' \) and \(-\text{O-p}' \) are true), then everything is permitted (\(\text{P-p}' \) and \(\text{Pp}' \) are true). And if there are people in an ideal world, which people? Does the product of a rape exist in an ideal world? Obviously not, Purtill claims. Secondly, the parenthetical remark in the passage of Hintikka's quoted above indicates that Hintikka believes norms may differ from world to world. But many cognitivists would argue that there is a single set of norms which applies to any possible world.

I do not find either objection to be particularly compelling. In Chapter IX I will express reservations about barring persons from ideal worlds. Lee Harvey Oswald may indeed exist in an ideal world, but some of his actions do not. It is actions, not persons, which make up the domains of ideal worlds. The progeny of rapes may well exist in

\textsuperscript{16}Hintikka [1971], p. 71.
\textsuperscript{17}Purtill [1973].
\textsuperscript{18}Purtill [1973], p. 5.
an ideal world and rapists (from the point of view of the actual world) may exist in ideal worlds, but rapes do not. As Purtill suggests, "we could, of course, begin tinkering with history, marry off the rapist and his victim in our deontic alternative, for instance."¹⁹ Purtill finds this undesirable since as we make all of the requisite alterations, the history of the ideal world becomes "more and more unlike the history of our world."²⁰ So be it. This is one way of finding out just how wicked our world happens to be.

Nor does it follow that everything is permitted in an unpopulated ideal world. In discussing the deontic analog of the Barcan formula, we saw that the truth values of atomic formulas for a given world are not determined if there are no individuals in that world. (If there are no persons, then there are no actions.) Hence, no truth values for '0(x)Ax' is specified by the semantic rules. Purtill is mistaken in supposing that '¬0(x)Ax' may be inferred from the supposition that there are no persons.

To the second objection, a cognitivist may easily suppose that moral principles are objective but the moral principles for any given world must reflect the morally relevant differences between that world and another world for which there is a different set of principles. For example, in a world in which people experience no pain, some prohibitions against some forms of torture could reasonably (and objectively)

¹⁹ Purtill [1973], pp. 3 and 4.
²⁰ Purtill [1973], p. 4.
be different.

Despite my rejections of Purtill's criticisms of the notion of an ideal world, his paper is not without merit. He offers arguments against Hintikka's characterization of

\[(40) \quad \Box(p \rightarrow q)\]

and

\[(41) \quad p \rightarrow \Box q\]

as expressing *prima facie* and actual obligations, respectively, and I concur with his argument against \((40)\) as a characterization of *prima facie* obligation, but I do not accept his reason for rejecting \((41)\) as a formalization of actual obligations. His argument in favor of the latter is very simply that if one is obliged to keep a promise (*i.e.*, '\(\Box q\)' is true), then it is trivially implied by any statement (*i.e.*, '\(p \rightarrow \Box q\)' is true). But there is an alternate way to make this same point: If '\(\Box q\)' is true, then the disjunction of '\(\Box q\)' with any other statement is also true. When Purtill's claim is rephrased in this manner, \((41)\) does not seem to be particularly objectionable. The alternative statement I have offered is certainly less ambiguous than the phrase "trivially implied by", and is much better justified by the truth table interpretation of the material implication sign.

I dislike \((41)\) as an expression of an actual obligation for a different reason. Hintikka's view that \((41)\) represents an actual obligation has always puzzled me because if you were to tell someone that they have an actual obligation to do something, you would most probably say 'You are under an actual obligation (this is the overriding obliga-
tion) to keep this promise, which is clearly not a conditional sentence and would be translated, quite simply, as 'Op'.

The treatment of (40) as representing prima facie obligations is another story. Suppose (40) represents the prima facie obligation to keep a promise. Suppose, furthermore, that murder for pleasure is absolutely forbidden, '0-r'. From '0-r', the prima facie obligation

\[ 0(r + q) \]

trivially follows. The promise 'p' will oblige you to keep it, 'q', 'no more or no less than any forbidden act'\(^21\) 'r' obliges you to keep it. Purtill concludes that there is no way in which HDL can distinguish between trivial prima facie obligations and important ones. Quite clearly, (40) cannot be used as a formal representation of prima facie obligation.

Does Purtill's excellent criticism require Hintikka to give up his criticism of Searle? I do not think so. In addition to the distinction between prima facie and actual obligations Hintikka has also given us the related distinction between deontic and logical consequences. They differ in that a deontic consequence, for example, has the same form as (40) but a deontic consequence is deontically valid. Insofar as Searle's second premise is "analytic" it may be viewed as either a logical consequence or a deontic consequence. Searle's argument is valid only if 'Promises ought to be kept,' is a logical consequence, i.e., 'If I make a promise, then I ought to keep it' is logically valid. Since there are some promises that may not be kept it is not

\(^21\) Purtill [1973], p. 6.
a logical consequence. Hintikka's argument is correct, it is his term-
inology that needs to be slightly modified.

I do not accept Purtill's ultimate conclusion that the founda-
tion laid by Hintikka will not support an adequate deontic logic.
On the contrary, HDL represents an extremely fruitful foundation for
further developments in deontic logic. But, quite obviously, the time
has come to either give up the syntactical structure of the standard
deontic logics (including HDL) or else find a new interpretation for
them. Since I have just said that HDL represents an adequate foundation
for further work in deontic logic, the reader may infer that I am about
to offer a new interpretation. But, be prepared for the introduction
of a syntax which will adequately handle \textit{prima facie} obligations
(Chapter IX).

In the previous chapter it was noted that we commonly say
things like

\begin{align*}
(43) & \text{This street ought to be widened.} \\
(44) & \text{There ought to be a chicken in every pot.} \\
(45) & \text{Everyone ought to be born with a sound mind and body.}
\end{align*}

and that these uses of 'ought' need not be construed as typically moral
uses of 'ought'; as claiming that there is a moral agent who is bound
by an obligation that he should be able to satisfy. If we suppose that
these do not represent claims about agents, then, according to Chapter
V, they may be taken as evaluative uses of 'ought' and, hence, still
part of normative discourse. The meaning of 'ought' in these sentences
is something like 'Would that the world were such that...' I shall
refer to these uses as *optatives* and I wish to propose an Optative Deontic Logic (ODL).

Optative uses are essentially evaluative uses of 'ought'. They are normative but not, in my view, moral uses. By that I mean that they are not directly related to moral agents and their actions, but, instead, indicate that some state of affairs is valuable, or preferable, or, to be precise, that a proposition is evaluated positively. A reasonable interpretation would read 'Op' as 'p ought to be true', and no restrictions on 'p' seem necessary since 'p' may even be analytic. It is not absurd, for example, to say 'Triangles have three sides and that's exactly the way things ought to be.' But it is absurd to say 'It ought to be that there is no evil and there is evil', since its formalization, 'O(p & ¬p)', is a deontic contradiction. The deontic logics which were introduced in the last chapter seem admirably suited to an optative interpretation. So I shall select the syntactic structure of either OT, OS₄, or OS₅ and informally interpret it in the manner suggested in this paragraph.

It was argued in the previous chapter that iterated operators posed a difficulty for standard deontic logics on a moral interpretation of 'ought'. On an optative interpretation, that difficulty also dissolves. 'OOp' means 'In all ideal worlds, p ought to be true.' The semantic understanding of 'POp' and 'OPp' and 'O00Op' are all clear even though stating them in English is awkward, e.g., 'POp', or something that we cannot imagine anyone ever wanting to say, e.g., 'O00Op'.
Both comments apply equally well to the formula \(((p + q) + r) + s + t) + u\).

The most obvious difference between OT, OS\(^4\), and OS\(^5\), however, is their treatment of iterated operators. Is there one that best coincides with an optative interpretation? I believe that there are grounds for rejecting OS\(^4\) and OS\(^5\). We might say of something that it ought to be (a street widened, for example), but negatively evaluate the fact that it ought to be (because the street ought to be widened given a larger population, more cars, and increasing pollution. So, even though the street ought to be widened, it is undesirable that it ought to be.) In ODL, this amounts to saying that the conjunction of 'Op' and 'O-Op' is consistent. This conjunction is consistent in OT but it is not consistent in OS\(^4\) and OS\(^5\). Hence, the proper syntactic structure for ODL is OT.

ODL may be quantified in the manner of HDL as long as the wffs are interpreted as evaluating actions either positively, 'O(\exists x)Ax', negatively, 'F(\exists x)Ax', or as being "all right", 'P(\exists x)Ax'. We must not interpret 'O(\exists x)Ax' as expressing an obligation, but as saying 'Would that the world were such that there is an instance of apple-picking' or 'It ought to be that there is an instance of apple-picking'.

The most important question to be asked about ODL is whether it avoids the paradoxes. Since ODL contains P, all of the "paradoxical" theorems are in ODL. However, their optative readings are not paradoxical.
The viciousness of Ross's paradox turned on the notion of satisfiability. This notion is not applicable to optative uses since optative oughts are essentially agentless and do not translate into duties. If (44), for example, is true, it does not follow that there is some agent who has the moral duty to put a chicken in every pot. When an optative ought obtains, there need not exist a moral duty or obligation to be satisfied. In the formula 'Op', 'p' is not a prescribed action when 'Op' is interpreted optatively. Hence, it does not seem non-sensical, but, on the contrary, a logical requirement that if 'p' is true in all ideal worlds, then 'p v q' is true in all ideal worlds.

An attempt to revive Ross's paradox for optatives might proceed as follows. Suppose that someone has the time, power, and inclination to produce an ideal world. By Ross's paradox, ideal worlds are filled with disjunctions, and one of the disjuncts, the "q" disjunct, may describe non-ideal actions. Our idealist may be misguided (by an Evil Genius, perhaps) into picking the "q" disjunct as the way to make these disjunctions true. This argument turns on the assumption that the "q" disjunct may be "non-ideal". If "q" is non-ideal, then it is either merely "all right" or "bad". If it is all right, then there is no problem. If it is bad, then 'q' partially describes all ideal worlds in question. In this case, if our idealist chooses 'q', the result will be an inconsistent world, hence, not an ideal world.

An optative reading of 'Pp' ('Well, I guess it's all right to widen the street.') helps to tell us whether or not the Good Samaritan paradox is a paradox of ODL. It is not. The theorem
means that if it is false that there is an ideal world in which Tom is robbed, then it is false that there is an ideal world in which it is true that Tom is robbed and Tom is given aid, and this is a deontic truth even if it is obligatory to give Tom aid. In this event, the man is given aid in the ideal world but not robbed. It might be objected that this is an odd ideal world in which men are needlessly given aid. Indeed, it is. But optative semantics suppose that 'p' and 'q' are logically independent propositions such as 'There is no evil' and 'April showers bring May flowers.' Clearly, if it is false that there is an ideal world in which there is no evil, then it is false that there is an ideal world in which there is no evil and in which April showers bring May flowers even if it ought to be that April showers bring May flowers. In the Good Samaritan case, however, there is a connection between Tom being robbed and giving him aid. We sense a deontic connection here, viz., it ought to be that if a person is robbed, then they are given aid (It is tempting to say we have a prima facie obligation to give aid to the victim of the robbery.). We certainly need not suppose that there is an absolute obligation to give aid to Tom. The obligation to give aid to Tom is dependent on his being robbed. So, in an ideal world, Tom is not robbed, it is not true that he is robbed and he is given aid.

The derived obligation paradox dissolves since the wffs of ODL are not interpreted as expressing obligations. 'Fp + 0(p \& q)' means nothing more than if 'p' is false in an ideal world, then either 'p' is
false or 'q' is true in that ideal world.

The remaining paradox, the contrary-to-duty paradox, will be treated in section 6.6.

6.5 Does Ought Imply Can?

Hintikka discusses another issue of traditional philosophical interest about which HDL has something to say, and that is the status of

\[(47) \quad Op + Mp\]

which has been interpreted as Kant's dictum that ought implies can.\(^{22}\) Hintikka denies that \((47)\) is the correct formalization of ought implies can. It is not, to use our new terminology, a logical consequence, but a deontic one. That is, \((47)\) is not a theorem, but

\[(48) \quad 0(Op + Mp)\]

is. The latter is a natural consequence of our semantics as the forest of Figure 10 indicates. Notice that in the ideal world, \(W_1\), 'p' is both obligatory and impossible. Since it is an ideal world, 'p' is true. If we suppose that what is necessary in an ideal world is also true in that world, then 'p' is false and the ideal world is an inconsistent world. Notice that we are not supposing that the ideal world is a full-fledged possible world. In other words, if \(W_1\) is a deontic alternative of \(W_0\) it need not also be an alethic alternative. If we did suppose this, then \((47)\) would also be a theorem. Inspect the forest for \((47)\) in Figure 11. Both 'Op' and 'L-p' require alternate worlds; one ideal

\(^{22}\)Anderson [1956], p. 183 and Prior [1958], p. 137.
Assume $\sqrt{-O(0p + Mp)}$

\begin{align*}
W_0 & \\
W_1 & \\
1 & \sqrt{-(0p + Mp)} \quad W_0-1 \\
2 & 0p \quad 1 \\
3 & \sqrt{-Mp} \quad 1 \\
4 & \sqrt{L-p} \quad 3 \\
5 & p \quad 2 \\
6 & -p \quad 4 \\
\end{align*}

\text{FIGURE 10}

Assume $\sqrt{-O(0p + Mp)}$

\begin{align*}
W_0 & \\
W_1 & \\
W_2 & \\
1 & p \quad W_0-2 \\
1 & -p \quad W_0-4 \\
\end{align*}

\text{FIGURE 11}
and one possible world. But need they be the same world? Or we might make the weaker assumption that 'p' is false in both worlds; that ideal worlds are full-fledged possible worlds. Whatever assumption is necessary to close the above tree, Hintikka says, "It nevertheless seems safe to say that no obvious and uncontroversial principle is forthcoming."\(^{23}\)

But, he also says, "The merits and demerits of such principles would require a longer discussion than can be undertaken here."\(^{24}\) A disappointment. However, I propose to partially fulfill the vacuum.

The question boils down to this: Is \(W_1\) (an ideal world) a possible world? The answer, I think, is both "yes" and "no". Let me explain. \(W_1\) is a logically possible world, but that is built into the semantic rules; every ideal world must be logically consistent, it must satisfy (C-). Is \(W_1\) a possible world in any other sense of 'possible'?

An affirmative answer would be presumptuous. Suppose one is using a combination of possible and ideal worlds to deal with a set of propositions about obligations plus one proposition to the effect that it is impossible for a moral agent to be in two places ten miles apart.

If that claim is formalized as 'L-p', then all of the possible worlds in which 'p' is false are physically possible worlds, and a deontically perfect world might not be a physically possible world. In fact, satisfying all of one's obligations often requires being ten miles apart at the same time, or some other physical impossibility. More often than not, if it is impossible to satisfy all of one's obligations, the impos-

\(^{23}\)Hintikka [1971], p. 84.

\(^{24}\)Ibid.
sibility is not a logical impossibility but a physical, economic, or, perhaps, a psychological one. In some cases, (47) might be a reasonable assumption to make, but not always. It must be assessed case by case on the basis of non-logical considerations.

The primary use of (47) is in situations where a person cannot, in some sense of "cannot", fulfill a duty. It might, for example, be psychologically impossible for someone to perform a duty even when convinced that he ought to do it, simply because he cannot bring himself to do it because of claustrophobia or some similar difficulty. In this case, (47) only plays a role if we deem the claustrophobia to be seriously debilitating; if, for example, the person chooses self-destruction over confinement in a small enclosure. But even in this kind of case, drawing the line is not a matter of logic. Physical impossibilities seem stronger but not sufficiently so. Otherwise, to avoid performing an onerous but obligatory task, a person could well place any number of physical or economic or psychological (hypnosis, perhaps) obstacles in the way, making it impossible to perform the required duty. It ought to be the case that when an obligation cannot be fulfilled, there is no obligation, but in this world claims such as 'I cannot make it to your mother's for the holidays with you and the children, dear, because I've hired some men to kidnap me and force me at gunpoint to spend the week with my lover," though true, are not mitigating. Furthermore, the strongest arguments in favor of an ought-implies-can principle interpret 'ought' as a personal imperative, not as an impersonal optative as we are now interpreting the 'O' operator. In fact, the examples of opta-
tives cited in section 6.4 suggest that optatives are often used when the argument of the operator is false and, perhaps, impossible in some sense. For example, one might wish that

\[(49) \quad 2 + 2 = 5\]
even though (49) is generally viewed as being inconsistent. On our semantic rules, however, it is only required that (49) be consistent with all other optatives. Along with (49), for example, one would have to wish that half of five is two.

Hintikka argues that our intuitions are not sufficiently sharp to keep (or, in general, recognize) the distinction between deontic and logical consequences clear in our moralizing. Kant, he says, had the same problem, but there is some textual evidence, Hintikka claims, that Kant intended the ought-implies-can principle to be a deontic consequence\(^{25}\). Hence the originator of the principle couched it in a view that requires it to be formalized as a deontic consequence. This distinction is clearly an important one, and it has fruitful and equally important consequences such as the criticism of Searle's attempt to derive ought from is.

Incidentally, at the risk of flogging a dead horse, I would like to point out that the result of this section deprives Anderson's logics of part of the evidence adduced in favor of them. (See section 4.4.)

\(^{25}\)See Hintikka [1971], p. 86 for the argument.
6.6 The Contrary-to-Duty Paradox

A bonus of adding quantifiers that Hintikka does not seem to realize\(^{26}\) is the contrary-to-duty paradox raised by Chisholm is nicely avoided. That paradox concerned a man obliged to go to the aid of his neighbors who also ought to tell them if he is coming and who, if he doesn't go, ought not tell them he is coming and who, in the end, fails to go to their aid. All of this is formalized by Chisholm as

\[
\begin{align*}
(50) & \quad 0p \\
(51) & \quad 0(p \land q) \\
(52) & \quad -p \lor 0-q \\
(53) & \quad -p
\end{align*}
\]

which according to Hintikka's semantics is an inconsistent set of statements. (See section 5.9 for the original discussion of this paradox.) The disturbing asymmetry of (51) and (52) seemed unavoidable when this paradox was discussed in Chapter V. With quantifiers, it can be avoided. Let 'Gx' and 'Tx' stand for the functions 'x is an instance of a man going to the aid of his neighbors' and 'x is an instance of a man telling his neighbors that he is coming to their aid,' respectively. The four original statements translate as

\[
\begin{align*}
(54) & \quad 0(\exists x)Gx \\
(55) & \quad 0(x)(Gx \land Tx) \\
(56) & \quad 0(x)(\neg Gx \land Tx) \\
(57) & \quad (x)\neg Gx.
\end{align*}
\]

\(^{26}\) Hintikka [1971], p. 102 where he says "...certain forms of contrary-to-duty obligations...require an altogether different sort of treatment."
Unlike the non-quantified formalization, this second set has the advantage of (55) and (56) being symmetrical yet neither one of them is implied by either (54) or (57), and the set is consistent. This leads me to believe that our original assessment of the paradox was wrong.

Perhaps the non-quantified version should have used

\[(58) \quad 0(-p + q)\]

instead of (52). But (58) was rejected, you may recall, because it was implied by (50). However, on the basis of Hintikka's interpretation (52) represents an actual obligation whereas (51) represents a *prima facie* obligation. The asymmetry is clearly untenable. There is no reason to suppose that the man's obligation to tell his neighbors that he is coming *is prima facie* but his obligation to not tell them he is coming is actual. It is more reasonable to suppose that they are both *prima facie* obligations. Hence, (58) should have been used originally.

Why wasn't it used? Primarily because one could also infer

\[(59) \quad 0(-p + q)\]

or anything else in the consequent position from (50). Is that bad? No, not if I am right about ideal world semantics and optative oughts. If it is true in an ideal world that a man aids his neighbors then it is also true in that same ideal world that if he does not aid his neighbors, then he does not tell them he is coming. It also follows that if he does not tell them, then he turns blue. But all of these inferences hold in the real world, too. They are no more mysterious or undesirable in ideal worlds than they are in the actual world. For this reason we should not accept Føllesdal and Hilpinen's diagnosis which was described
in section 5.9. They argue that (52) must be a claim about "almost" ideal worlds since in those worlds 'p' is false whereas 'p' is true in all ideal worlds, and, consequently, another operator is necessary. I believe that the quantified formalization shows this to be unnecessary and on an ideal world semantics that (52) should be replaced with (48) in a non-quantified formalization. The oddness of this suggestion comes from the fact that we are inclined to view the original four statements as claims about the duties of the man who ought to go to the aid of his neighbors rather than claims about how the world ought to be.

It would be quite nice, however, to be able to derive what one ought to do when one fails to keep one's obligations. But this requires some means of formalizing those statements that look like

\[(60) \quad \text{You, Al, have an obligation to pay this debt.}\]

and

\[(61) \quad \text{You, Al, ought to pay this debt.}\]

These cannot be completely captured in HDL since they claim that an agent has an obligation which the agent should satisfy. Formalizing claims of this sort as 'Op' gives rise to Ross's paradox. The rest of this paper is devoted to a search for a way of formalizing (60) and (61) which avoids Ross's paradox. It begins by looking at, and rejecting, a proposal made by H. N. Castaneda. Castaneda takes the other quantificational route of quantifying over agents. The predicate function 'Ax' is read as 'The agent, x, does A.' At the outset we may apply Hintikka's arguments in favor of quantifiers and point out that whether 'A' is interpreted as an act individual or as an act type, the calculus is
going to be much more limited than if it also permitted quantification over act individuals. This comment will not be repeated in the next chapter. Instead, the last chapter develops the best of all possible worlds; a deontic logic which quantifies over agents as well as acts.
7.1 Imperatives and Obligation

Castaneda begins his most recent exposition\(^1\) with a distinction which is, I hope to show, extremely fruitful for deontic logic. It is the distinction between the "ought-to-do" and the "ought-to-be", as Castaneda calls them. He suggests that both of these notions are equally deserving of logical investigation. Since the standard deontic logics deal with the ought-to-be, Castaneda will fill the void by developing the logic for the ought-to-do. The only other writer who has consciously drawn the distinction is von Wright who did so in a paper read to a meeting of the Creighton Club (New York State Philosophical Association) on October 27, 1972\(^2\). He suggested that the logic of ought-to-be is so similar to the logic of traditional modal logic that it does not represent an autonomous branch of logic. The previous chapter suggests that this claim, if not false, is at least misleading. It may well be that the logic of ought-to-be does not represent an autonomous branch of formal logic, but it certainly warrants a separate chapter in the area of philosophical logic. The deontic logic which von Wright

\(\text{\footnotesize\textsuperscript{1}}\) Castaneda [1970].

\(\text{\footnotesize\textsuperscript{2}}\) von Wright [1973].

162
presented in this recent paper is not nearly as interesting as the action logic on which it is built. His action logic plays an important role in the next chapter.

Castaneda points out that in the standard systems "The idea of who is to realize the obligation is not considered, so that the approach can handle very nicely genuinely impersonal statements like 'There ought to be no pain', meant merely to articulate something about the universe." He concludes: "In short, deontic statements divide neatly into: (i) those that involve agents and actions and support imperatives, and (ii) those that involve states of affairs and are agentless and have by themselves nothing to do with imperatives." As we shall see, the arguments of deontic operators in Castaneda's deontic logic will be imperatives instead of the propositions employed by standard deontic logics. Consider the following three sentences:

1. John is going to the store.
2. John, go to the store!
3. It is obligatory that John go to the store.

and notice that (3) is easily formed from (2) by dropping the comma and the exclamation point, and preceeding the entire expression with 'It is obligatory that...'. These three sentences are said to correspond with one another. For Castaneda, correspondence lies in the fact that all three sentences involve the same action. The three sentences all

---

³ Castaneda [1970], p. 450.
⁴ Castaneda [1970], p. 452.
consider, to use Castaneda's terminology, the same action, but they consider it in different ways. He illustrates this point with the following example:

(4) It is obligatory both that today Jones sign the green card if he comes late, and that he works after 5:00 pm if he-signs-the-green-card, \(^{5}\)

and points out that the emphasized phrase is considered prescriptively, and the hyphenated phrase is considered as a condition and as such it is considered as having a truth value whereas the prescriptively considered phrase is not thought to have a truth value. Actions prescriptively considered are like statements of actions in that (i) they can be parts of larger statements, (ii) they stand in entailment relations, and (iii) they can (only in part) be the objects of cognition. They are unlike statements in that they are (a) the subjects of deontic properties, (b) do not have truth values, and (c) cannot be the whole objects of purely cognitive attitudes like belief or supposition. Hence, Castaneda argues, deontic operators apply to actions, not to statements. He says, "Doubtless, a prescription is closely related to its corresponding performance statements, and may be supplanted by one of them as a constituent of a deontic assertion." \(^{6}\) But it must be done in the following way:

(5) John ought to make 'John goes to the store' true, which is quite different from the standard treatment which produces

\(^{5}\)Castaneda [1967], p. 30.

\(^{6}\)Castaneda [1967], p. 31.
(6) It is obligatory that John goes to the store, and the sense of (6) is somewhat more impersonal than (5). In (5) it is quite clear that John is the party responsible for seeing to it that John goes to the store, whereas in (6) it is not clear. It could be the responsibility of John's parents, teacher, guardian angel, or, perhaps, even God, with John only being obliged to obey those in authority. Before going on, I would like to exhibit a locution that even more perspicuously performs the function I have been attributing to (5), and that is

(7) You, John, ought to go to the store.

This last locution is as clear a case of a prescriptive use of 'ought' that one can find. There is no doubt that the speaker is recommending, exhorting, or whatever, that John go to the store. (6), on the other hand, which is the paradigm use of 'ought' explicated by standard deontic logics, suggests nothing about what Castaneda calls the agency of John, or anyone else, for that matter. (6) does not tell us who is bound by the obligation—the agent whom the obligation compels to act. And (7) differs from an imperative in that (7) tells John what he should do, but the imperative 'Go to the store, John!' tells John what to do, and such things as war crime trials suggest that an imperative can be in force which orders a forbidden action to be done.

The notion of agency presents itself quite nicely in these two pairs of statements:

(8) Paul hit Mary.

(9) Mary was hit by Paul.
Traditionally, the statements which compose the first pair do not differ in meaning, and choosing either the active voice or the passive voice is a matter of style, not understanding. Things are much different, however with the second pair. In this case, the transformation from the active voice to the passive voice involves at least a loss of agency and perhaps a shift in agency. It also clearly results in a change of meaning. In the former, Paul is the mandated agent, and in the second it seems that Mary is the mandated agent. Now consider a third pair:

(12) Paul ought to hit Mary,
(13) Mary ought to be hit by Paul,

and notice that we are again faced with a change of agency, at least the hint of one. In (12), if anyone is claimed to have an obligation, it is Paul. (13), however, makes no suggestion that Paul has an obligation, and if it is claiming that anyone does, Mary is the only plausible candidate. The reason that the shift in agency in this pair is very weak (if there is a shift at all) is that (12) and (13) are closer to being examples of the ought-to-be sense of 'ought'. Hence agency is an important difference between the ought-to-do and the ought-to-be. The ought-to-do involves actions that are prescriptively considered; actions for which some agent is responsible for seeing that they occur or do not occur. For Castaneda, this means that the arguments of deontic operators which represent the ought-to-do must be prescriptives.
In my opinion, Castaneda has undeniably made it clear that agency is a central feature of the ought-to-do. He has not made as convincing a case that prescriptives are the arguments of deontic operators of the ought-to-do sort. I say this because the distinction between prescriptives and imperatives tends to be blurred in Castaneda's discussions of them. In many places, Castaneda uses the terms interchangeably. Furthermore, prescriptives and imperatives must be distinguished from mandates. The term 'mandates', he says, refers to a class of entities which consists of commands, orders, requests, pieces of advice, petitions, and entreaties, and any of these may be expressed by "one and the same imperative sentence." Apparently, an imperative is a grammatical structure of a natural language and mandates represent the various uses of this grammatical structure. A prescriptive, however, is a logical entity which is expressed in natural languages by an imperative sentence. If my interpretation of Castaneda is correct (textual evidence is quite scarce), then the relationship between imperatives and prescriptives is analogous to the relationship between indicative sentences and propositions. And I suppose that mandates would correspond to asserting, informing, affirming, claiming, stating, and, perhaps, supporting, concluding, and reminding. Hence, any given prescriptive may be interpreted as a command or an order or any other type of mandate, and that prescriptive will usually be expressed in ordinary language with an imperative. As best as I can tell, this is

7Castaneda [1970], p. 452.
how Castaneda is using these terms.

What all of this brings to mind is Hare's neustic-phrastic distinction. Hare believes that

(14) Shut the door!

and

(15) You are going to shut the door, have something in common, viz., your shutting the door. This he calls the phrastic. They differ in that (14) results from combining this phrastic with an imperative neustic, and (15) is an application of an indicative neustic to this same phrastic. Hare employed this device to distinguish imperatives and indicatives, but it can be extended, it seems to me, to include an interrogative neustic which combined with the phrastic mentioned above yields 'Are you going to shut the door?'

Furthermore, a deontic neustic seems quite plausible: 'You are obliged to shut the door.' This suggests that one way to account for agency is to hold that ought statements and imperatives are both built up from a common element (Hare's phrastic) and that, unlike indicative and interrogative neustics, imperative and deontic neustics just happen to have the feature of agency in common, rather than suppose, as Castaneda does, that ought statements have imperatives as components. Castaneda does not rule out this possibility (in fact, he doesn't consider it at all). On the contrary, I shall present a deontic logic which contains what amount to indicative phrastics and agential phrastics. I will show that deontic statements can be formed from both types of phrastics;
the ought-to-be is formed from indicative phrastics and the ought-to-do is formed from an agential phrastic. Furthermore, imperatives will be formed from agential phrastics independently of the ought-to-do, but I believe that there is a third type of deontic statement which can be formed from an imperative. These I shall call prescriptives.

7.2 The Formal Structure of CDL

How, then, does the logic of prescriptives differ from the logic of indicatives? As it turns out, there is no difference, according to Castaneda. The syntactical elements of CDL (Castaneda's deontic logic) are those of predicate logic plus the familiar deontic operators, 'O', 'P', and 'F'. Added to these is an underline which is used in predicate functions, e.g., 'Z(x,y)' to indicate the agency of 'x', so that 'Z(x,y)' might be read as the imperative 'x, do Z to y'! In the formation rules listed below, 'p', 'q', and 'r' range over indicatives (propositions or truth functional compounds of propositions). 'A', 'B', and 'C' range over prescriptives, and 'p*' and 'q*' range over both indicatives and prescriptives.

There are two types of wffs. Indicatives are sequences of syntactical elements having one of the following forms:

(FR1) \[ Z(x_1,\ldots,x_k,\ldots,x_n) \] where \( Z \) is an \( n \)-adic predicate and each \( x_i \) is an individual constant and variable, and no \( x_j \) is underlined.

(FR2) \[ (\neg p) \]
Prescriptives are sequences of syntactical elements having one of the following forms:

(FR6) \[ Z(x_1, \ldots, x_i, \ldots, x_n) \] where \( Z \) is an \( n \)-adic predicate, each \( x_i \) is an individual variable or constant, and at least one \( x_i \) is underlined to indicate the agency of \( x_i \).

I have no idea why (FR6) specifies "at least" one agent term be underlined rather than "exactly" one. "Exactly" one seems quite sufficient. I suppose we could interpret \( Z(x, y) \) as 'x and y, dance together!' But 'x and y, dance together!' seems to entail 'x, dance with y!' and 'y, dance with x!' Unfortunately, this logical relationship does not hold in CDL; \( Z(x, y) \) does not imply 'x, dance with y!' and 'y, dance with x!' unless a special assumption is introduced. If Castaneda specifies "exactly" one instead of "at least" one, assumptions of this sort are not necessary.

In addition, CDL employs the usual definitions of 'P', 'F', 'v', '→', and '≡'. The axioms of CDL are:

(A1) \( p^* \), if \( p^* \) has the form of a truth-table tautology.
(A2) \[ OP \rightarrow PA \]
(A3) \[ (OA \& OB) \rightarrow O(A \& B) \]
(A4) \[ (p \& OA) \rightarrow O(p \& A) \]

The rules of inference are:

(MP) If \( p^* \) is a theorem and \( p^* \rightarrow q^* \) is a theorem, then \( q^* \) is a theorem.

(DR1) If \( A \rightarrow B \) is a theorem and contains no quantifiers, then \( OA \rightarrow OB \) is a theorem.

(DR2) If \( p \rightarrow A \) is a theorem and contains no quantifiers, then \( p \rightarrow OA \) is a theorem.

(UG) If \( p^* \) is a theorem, then \((x)p^* \) is a theorem.

In section 7.4 I will argue that CDL does not adequately fulfill Castaneda's intentions. In the meantime, I should point out that my presentation simplifies the original. Castaneda actually presented a series of deontic logics such that the obligations of one take precedence over another. One of these logics represents the logic of the "pure overriding ought". This logic has axioms which differ slightly from the others, most notably

(16) \[ OA \rightarrow A \]

This is Castaneda's device for distinguishing _prima facie_ and absolute obligations. (16) is an axiom of the logic of absolute obligations, so absolute obligations imply their corresponding imperative. I don't think that they do, and I also think that Hintikka's method of distinguishing _prima facie_ and absolute obligations is superior to Castaneda's.
Though wrong (see section 6.4), Hintikka's method at least has a device for indicating the conditions upon which the *prima facie* obligations depend. Furthermore, it seems to me that I may inform you of any number of obligations that you in fact have yet not tell you to satisfy them. I may even urge, recommend, advise, counsel, and implore you to do something by saying "You really ought to repay that loan," and not be telling you to repay the loan. This becomes especially clear in situations where the advisor also has command authority over the advisee. A sergeant, for example, might tell a new recruit that he is such a misfit that he ought to commit suicide, yet isn't clear that he has told or ordered the recruit to do so. Further evidence for this is that there would be a clear difference in the recruit's third person account. Notice the difference between 'He told me to do it' and 'He told me I should do it.' It does not seem to me that either one implies the other.

7.3 The Semantics of CDL

Every atomic formula of CDL is assigned either the semantic value '1' or the semantic value '2'\(^8\), and these same semantic values are assigned to molecular formulas in the following manner: '1' is assigned in exactly the same manner that 'true' is assigned by propositional logic, and '2' is assigned in exactly the same manner that 'false' is assigned by propositional logic. Note that 'Z(x,y)\(^1\)' and 'Z(x,y)\(^2\)' and

---

\(^8\)The use of numerical values is explained by the fact that the atomic formula 'A' has no truth value.
'Z(x,y)' and 'Z(x,y)' need not be assigned the same semantic value. Consequently, CDL without quantifiers and without deontic operators is nothing more than propositional logic. Adding quantifiers produces the first order predicate calculus. Castaneda did this deliberately. His reasons are examined in the next section.

There are two semantic rules which distinguish CDL from predicate logic. The first is

(17) If there is a world in which p (an indicative) is true, then p is true in every world.

Castaneda's rationale for (17) is this:

Thus a piece of rule-making confronts the real world, with all its facts and laws of nature, with other possible practical worlds in which the same facts and laws of nature hold, but different decisions and orders take place: the making of a set of rules is, at bottom, nothing more (on this analysis) than the adoption of a system $S_B$ of alternative practical worlds which share both the same facts and the same set $\beta$ of prescriptions or commands.9

I can only say that I don't find the sharing of facts by practical worlds to be an important feature of deontic concepts. In fact, I would think that rule-makers would be more interested in providing a set of rules applicable in a world no matter what the facts are. But this is not an important issue; CDL has more serious difficulties.

The other rule peculiar to the semantics of CDL is:

(18) OA is true in the original world if and only if A is true in every other world.

It is not immediately apparent, but (18) implies a derived semantic rule:

(19) OA is false in the original world if and only if there is some world in which A is false.

As before, a formula is CDL-valid if and only if its negation has no consistent model. For no reason known to me, Castaneda does not list semantic rules for quantifiers. I believe it is safe to say that the standard interpretation of quantifiers applies to CDL.

Given these rules, CDL, as Castaneda claims, is quite non-standard. The theorems of CDL include some familiar faces as well as some new ones:

\[(20)\quad O(A \land B) \equiv (OA \land OB)\]

\[(21)\quad P(A \lor B) \equiv (PA \lor PB)\]

\[(22)\quad O(p \lor A) \equiv (p \lor OA)\]

\[(23)\quad O(p \land A) \equiv (p \land OA)\]

\[(24)\quad P(p \lor A) \equiv (p \lor PA)\]

\[(25)\quad P(p \land A) \equiv (p \land PA)\]

The last four rules indicate that an indicative which falls within the scope of a deontic operator may always be placed outside the scope of that operator. So, it should be no surprise to discover that

\[(26)\quad O(A \rightarrow p) \equiv (OA \rightarrow p)\]

and

\[(27)\quad O(p \rightarrow A) \equiv (p \rightarrow OA)\]

are also theorems. This means that Castaneda does not recognize the distinction introduced by Hintikka between deontic and logical consequences. This is somewhat surprising in view of Castaneda's claim that
"A tremendously valuable distinction Hintikka makes is that between logical consequence and deontic consequence."\(^{10}\) In addition, CDL counts

\[(28)\quad (0(p \rightarrow A) \& p) \rightarrow OA\]

among its theorems. Hintikka, however, does not think that (28) is logically true because of the aforementioned distinction.\(^{11}\) Castaneda, it seems to me, owes us an explanation.

Castaneda's semantic rules are not sufficient to account for all of the theorems that can be syntactically derived. Suppose that axiom (A2) is false. Then 'OA' is true and '0-A' is false, and, according to semantic rule (18), 'A' is true in all other worlds and 'A' is false in all other worlds. But this is a contradiction, hence, the original world is inconsistent and (A2) is CDL-valid. The semantic rules for the standard deontic logics would produce the same results. But if we suppose that there are no deontic alternatives, which the semantic rules for CDL apparently allow us to do, then (A2) has a consistent model, hence (A2) is not CDL-valid. The standard deontic logics do not have this problem. The semantic condition (C.P) does not allow us to suppose that the model for (A2) has no deontic alternatives. Clearly, Castaneda is in need of a semantic condition similar to (C.P).

The most serious defect of CDL, however, is our old friend Ross's paradox. It turns out that 'OA \rightarrow 0(A \lor B)' is a theorem of CDL. In the previous two chapters I have argued that the difficulties of

\(^{10}\) Castaneda [1970], p. 450.

\(^{11}\) See the discussion of this in section 6.3.
Ross's paradox cannot be avoided if 'OA' is interpreted to mean that an
agent is bound by an obligation which that agent should be able to ful-
fill. All of the examples that have been given to support the paradox
interpret 'OA' as expressing a duty or an obligation. I have argued
that Ross's paradox can only be avoided by interpreting 'OA' in some
other way, and that the optative interpretation successfully avoids the
paradox. It seems quite clear to me at this point that any interpreta-
tion of 'OA' which involves the notion of agency requires that 'OA →
0(A v B)' not be a theorem. I believe that this and all of the other
difficulties of CDL stem from an inadequate treatment of agency. This
is the task taken up in the next section.

7.4 A Critique of CDL

Perhaps the most striking feature of the syntax of CDL is
that prescriptions may be treated as components of truth functions. This
syntactical feature is echoed in the semantics of CDL since the semantic
values are assigned in the very same way that truth values are assigned.
Castaneda does this because he believes that the logic of imperatives is
just like the logic of propositions. He defends this view as follows.
First, both of the following arguments seem valid:

(29) X, do A!
     Therefore: X, don't fail to do A!

(30) X, don't fail to do A!
     Therefore: X, do A!

which he believes establish the law of excluded middle for imperatives.
Next, he claims that all of the following rules are intuitively valid:

(31) 'X, do A' implies 'X, do A and do A',

(32) 'X, do A and B' implies 'X, do A',

(33) 'X, don't both do A and fail to do B' implies 'X, don't do both of the following: one, fail to do B and some action C; two, do C and A.'

Since these rules obviously parallel the axioms of Rosser's basis for propositional logic, and since

(34) 'X, do A' and 'X, don't do both: A and not-B' imply 'X, do B',

is parallel to *modus ponens* and is also intuitively valid, it should be clear that the logic of imperatives is identical to the logic of propositions. Imperative inferences may be evaluated by using the standard techniques of propositional logic. Simply allow the propositional variables to stand for imperatives and apply all of the usual rules of propositional logic. Castaneda does not introduce a "prescriptive" operator which may have other operators within its scope. All compound prescriptives are simply truth functions of prescriptives, and the logical laws of prescriptives are co-extensive with the laws of propositional logic. I will subsequently argue that Imperative inferences are more complex than this analysis suggests. In the meantime, I am not entirely happy about the idea of allowing prescriptives to freely replace propositional variables in wffs of propositional logic.

In CDL, not only is

(35) \( A \rightarrow (B \rightarrow A) \)
well-formed, but, by axiom (A1), it is a theorem as well. It would
appear that

(36) If X, do A, then if Y, do B, then X, do A,
is a permissible interpretation of (35). But (36) is obviously not
well-formed in English. An obvious retort, which, as we shall see,
Castaneda makes, is that '→' is not to be read as 'if...then' when
flanked by prescriptives. 'A → B' must be read as either 'X, don't
both do A and not-B' or as 'A implies B'. This means that '→' has two
readings depending on whether it connects indicatives or prescriptives.
What does one do in mixed cases? The expression 'p → A' makes perfect
sense: 'If p is the case, then X, do A', but how does one make sense
of 'A → p'? The only reading that makes sense is 'A implies p'. But
how can a prescriptive imply an indicative?

I don't believe that 'X, do A' interpreted non-propositionally
implies anything. But it might be said, suppose that 'X, do A' repre-
tsents the command 'You, go to the store!' and that it has the semantic
value '1'. Then 'X, do A' implies the indicative sentence that the per-
son who issued the command had the authority to do so. In my opinion,
this follows not from 'X, do A' but from 'X, do A has the semantic value
1', which, unlike 'X, do A', is a proposition. 'X, do A' implies that
the person who gave the command had the authority to do so if and only if
'X, do A' is "true" or "in force" or "felicitous" or however '1' is
interpreted. According to Tarski's criterion, whenever we speak of one
proposition implying another we may just as easily speak of the fact
that one proposition is true implies that another proposition is true.
If Tarski's criterion has a prescriptive counter-part, then 'X, do A' has the semantic value of '1' if and only if X, do A. Castaneda's semantics certainly seem to require this. Hence, wherever 'X, do A' occurs we may replace it with 'X, do A has the semantic value 1', and, consequently, 'X, do A' may be interpreted propositionally. Thus, I see no difficulty in agreeing with Castaneda that "The grammatical fact that 'if' does not govern imperative sentences shows, not that 'conditional' imperative tautologies do not correspond to valid imperative inferences, but only that such tautologies are not expressed by means of the word 'if'."12

Still, I cannot help but wonder whether that grammatical difference signals a logical difference. Especially since the grammatical differences run deep. All imperatives are second person; propositions need not be. Unlike a proposition, an imperative most naturally appears as the consequent of a conditional, not as an antecedent. It is not clear how truth-functional negation is to be applied to an imperative because one may "negate" in a broad sense of the word, an Imperative in one of two ways. First, the "opposite" of

(37)  Do this!

seems to be

(38)  Don't do this!

But the "denial" of (37) would have to be something like

(39)  It is not the case that 'Do this!' is in force.

It is not clear which of these the formula '-A' of CDL is to represent.

---

First, the readings contained in (29) and (30) suggest that (38) is the correct interpretation of '¬'. But this interpretation together with the semantics of CDL requires that every action or its negation is prescribed. Surely, an action can be such that neither it nor its negation is prescribed. This seriously jeopardizes Castaneda's two-valued analysis of prescriptives. It seems to me that (39) is the interpretation of '¬' which is most amenable to a two-valued analysis of prescriptives. But (39) also has the effect of turning a prescriptive into something that looks very much like an indicative. This must be avoided if 'O' is to apply to both 'A' and '¬A' in CDL. It should be clear, however, that no matter which interpretation of '¬' Castaneda chooses, CDL has no device for formalizing the other sense of negation. For example, suppose 'A' represents the command 'Sit there!' and '¬A' represents 'The command "Sit there!" is not in force'. Then, the command 'Don't sit there!' cannot be formalized in terms of 'A'. This is unfortunate since

(40) 'Don't sit there!' implies that the command 'Sit there!' is not in force.

appears to be a logical truth of commands. CDL can not even formalize it. I believe that the difference between (38) and (39) is worth formalizing. Second order ordinary command language, however, speaks not of commands being "negated" or even "denied". Instead, they are "overridden", "withdrawn", "countermanded", or simply non-existent, and the sort of relation to a command which withdrawal represents is not the same as that of a countermand. The difference between a withdrawn command and a countermand is roughly similar to the difference between (38) and (39)
although withdrawing a command implies the prior existence of a command whereas (39) does not. A countermand also suggests that there was once a different command in force whereas (38) implies no such thing.

The failure of CDL to adequately deal with prescriptive negation is the result of not seeing that the logical behavior of agency is more complex than CDL indicates. Agency, as we are using the term, and this use coincides with Castaneda's, refers to the fact that in addition to simply doing an action, an agent can be responsible for an action, that the agent is mandated, as Castaneda would say. An agent can either turn on a light when entering a room or not. These two alternatives are mutually exclusive and jointly exhaustive. But an agent can take the responsibility (or be ordered, or have promised) for seeing to it that the lights in a room are on or take the responsibility for seeing to it that the lights are not on. But between these two alternatives there are two alternatives wherein the agent is neither responsible for the lights being on nor for the lights being off. We might represent these four possibilities as

\[(41) \text{ } X \text{ sees to it that } X \text{ does } Z,\]
\[(42) \text{ } X \text{ sees to it that } X \text{ does not do } Z,\]
\[(43) \text{ It is false that } X \text{ sees to it that } X \text{ does } Z,\]
\[(44) \text{ It is false that } X \text{ sees to it that } X \text{ does not do } Z.\]

Notice that the locution 'X sees to it that X does not do A' is stronger than 'X does A'. The latter can be true even if X is unconscious or has never considered doing either 'A' or 'not-A', but under these conditions the former cannot be true.
Suppose that

\[(45) \quad x[Z(x)],\]
\[(46) \quad x[-Z(x)],\]
\[(47) \quad -x[Z(x)],\]
\[(48) \quad -x[-Z(x)],\]

represent (41), (42), (43), and (44), respectively. Then the two types of prescriptive negation could be represented as

\[(49) \quad 1x[-Z(x)]\]

and

\[(50) \quad -1x[Z(x)],\]

which represent 'Don't do this!' and 'It is not the case that Do this! is in force', respectively. Notice that once an agent logic is available which can capture the difference between commission, represented by (45), omission, formalized by (46), and neither committing nor omitting, then a prescriptive logic which is capable of formalizing both sorts of negation is a natural extension. But another natural extension is a deontic logic. So that, instead of appending an exclamation point to agent formulas (e.g., (45) and (46)), a deontic operator produces deontic formulas, e.g.,

\[(51) \quad 0x[Z(x)],\]
\[(52) \quad 0x[-Z(x)],\]
\[(53) \quad -0x[Z(x)],\]
\[(54) \quad -0x[-Z(x)].\]

This, I believe is evidence in favor of an alternative not considered by Castaneda (see section 7.1), namely, that prescriptives and deontic
formulas are both constructed from agential phrastics rather than
deontic formulas being constructed from prescriptives. A logic of
agency, based on these suggestions and which, as it must, avoids Ross's
paradox is constructed in the next two chapters. In the meantime,
足够的 has been said, I believe, to show that CDL does not adequately
capture the logical behavior of agency.

Among the theorems of CDL we find

(55) \[ A \rightarrow (A \lor B). \]

From the prescriptive 'A', the prescriptive 'A \lor B' follows. Let us
suppose that the particular mandates being represented are orders.
The point of orders is that they are to be obeyed. Let 'A' be 'Stand
at attention!', and let 'B' be 'Kill your superior!' Then, by (55),'stand at attention!' implies 'Either stand at attention! or Kill your
superior!' So, Private Concern, being an obedient soldier and exercis­
ing what clearly seems to be his choice, kills his superior, then stands
at attention awaiting further orders quite confident that he has obeyed
all outstanding orders, even some of the derived ones, for good measure.
I believe that it may be safely said that we are in a position to assure
Private Concern that his defense that he was only following orders will
not result in an acquittal at his court-martial.

Castaneda might respond that the connective in 'A \lor B' is only
a truth function and no one familiar with truth tables can doubt that
if Concern is ordered to stand at attention, then Concern is ordered to
stand at attention or Concern is ordered to kill his superior. In fact,
I believe that this is the only way that 'A \lor B' should be interpreted
as a wff of CDL. But this move treats 'A' and 'B' as propositions of no particular sort and there is no longer any reason for distinguishing between prescriptives and indicatives with any formal device. The problem is that the logical structure of imperatives is much more complex than CDL would lead one to believe.

To further demonstrate this point, I will attempt to show that the logical behavior of

(56) You, fix it or repair it!

need not be the same as

(57) Either you, fix it! or You, repair it!

Yet CDL is incapable of expressing that difference. That there is a difference between them can be seen by comparing

(58) You, Al, fix it or replace it!

with

(59) You, Al, fix it! or You, Ben, replace it!

which must be translated as

(60) \(G(a) \lor H(a)\)

and

(61) \(G(a) \lor H(b)\),

respectively. It should be clear, first of all, that (59) does not imply

(62) You, Ben, replace it if Al fails to fix it!

because of the truth functional interpretation demanded of (59). This means that from (59) it is possible to infer 'You, Ben, replace it!' only if '¬G(a)' is the case, and '¬G(a)' is interpreted by Castaneda as
'You, Al, don't fix it!' (62) is formalized in CDL as

\[(63) \quad -G(a) \rightarrow H(b)\]

and in virtue of the difference in underlining between (63) and (61) it should be clear that (61) does not logically imply (63). Since (59) does not seem to imply (62), this is as it should be.

One might argue that (59) does imply (62). I think this implication requires interpreting (59) as something stronger than what is expressed by (63)—as saying that Al and Ben are jointly responsible for seeing to it that either Al fixes it or Ben replaces it. The best CDL can do to represent this is

\[(64) \quad (G(a) \lor H(a)) \land (G(b) \lor H(b)).\]

It should be evident that (64) does not logically imply (63) in CDL. Notice the underlining in (63) and (64) and recall that 'G(a)' and 'G(a)' are logically independent. Hence, if (59) is interpreted in such a way that (62) does follow, CDL is not able to show that it does. This is not as it should be.

Furthermore, (58) as it stands implies

\[(65) \quad \text{You, Al, if you fail to fix it, replace it!}\]

which translates into CDL as

\[(66) \quad -G(a) \rightarrow H(a).\]

Again, notice the underlining. And notice that in virtue of the underlining that (60) does not, according to CDL, imply (66). This, too, is not as it should be.

Formally, CDL is not rich enough to capture the difference between (58) and (59). The main reason for this deficiency is that it
lacks the formal machinery for dealing with the complexities of negation and agency. There is an important difference between

(67) Al wears a red shirt,

and

(68) Al sees to it that Al wears a red shirt.

Agency, at least the sort of agency needed for an adequate deontic logic, is not simply a matter of a person performing some action. A person may well perform an action and not be responsible for the action. We need to be able to distinguish between the mere performance of an action and being responsible for the performance of an action. Consider the difference between (68) and

(69) Al's mother sees to it that Al wears a red shirt.

The difference between (68) and (69) is that different agents are responsible for (67) being true, and this is a difference which cannot be expressed by (67) alone, and it is a difference which an imperative logic must be able to express to adequately deal with the inference from (58) to (65) and the assumption which allows the move from (59) to (62).

It is a disadvantage of CDL that this difference cannot be formalized. It is not difficult to see that the difference is an important one for deontic logic. The statement

(70) Al ought to wear a red shirt,

does not contain a confident and clear specification of who is responsible or obliged to see to it that Al wears a red shirt. In the absence of any clear indication we tend to assume it is Al. The agency is made clear, however, and the agent is different in
(71) You, Al, ought to see to it that you wear a red shirt. 
and 
(72) You, Ben, ought to see to it that Al wears a red shirt. 
While (67) is a performance statement related to (71) and (72), it is 
not, to use Castenada's nomenclature, the corresponding performance 
statement to (71) and (72). The performance statements which correspond 
to (71) and (72) are 
(73) Al sees to it that Al wears a red shirt, 
and 
(74) Ben sees to it that Al wears a red shirt, 
and they both imply (67), but CDL cannot formalize these logical rela-
tionships without making what are, in the formalism of CDL, non-logical 
assumptions. The next two chapters develop an improved deontic logic 
which is based on a combination of ODL and an improved logic of agency.
8.1 The Uses of 'Ought'

Earlier, I introduced the distinction between the optative sense of 'ought' and the prescriptive sense. That distinction is in need of a slight bit of refining. The prescriptive sense is somewhat wider than I have heretofore suggested. For reasons that will shortly become clear, I will speak of the optative sense and the non-optative sense. These two senses are not to be construed as corresponding to different types of obligations that a person may have. A person may have both legal obligations and moral obligations, for example, but there are no optative obligations. All moral obligations, I want to suggest, are expressed by the non-optative sense of 'ought'. And the optative and the non-optative senses refer to separate logical uses of the word 'ought'. That is the thesis which this chapter will defend.

So far, this distinction has remained at an intuitive level, yet the evidence in favor of it has been growing. First, it was argued that the sense of 'ought' in which people are held to have obligations which they may or may not fulfill, cannot provide a meaningful interpretation of iterated operators because 'Op' is not the kind of thing which an agent, no matter how talented, can do to satisfy the obligation
expressed by 'OOp'. Hence, 'OOp' cannot express the sort of obligation that an agent can have. But, if we interpret 'Op' as 'p is true in all ideal worlds', then 'OOp' can be handled quite nicely, viz., 'Op' is true in all ideal worlds. And I have suggested that interpreting 'Op' as 'p is true in all ideal worlds' is an intuitively reasonable way of interpreting evaluative claims. On this interpretation, ODL becomes the criterion for consistency of evaluations: A set of evaluations is consistent if and only if the statements being positively evaluated are themselves consistent, and whatever is evaluated as "all right" or "neither good nor bad" (i.e., permitted) must be consistent with all statements which have been positively evaluated. This is just a restatement of the semantic basis of ODL.

Secondly, Ross's paradox makes it impossible to interpret 'Op' as 'S has an obligation to do A' for the familiar reason that S then has the obligation to do 'A or B' in addition to the obligation to do A. Once again, we cannot make sense out of satisfying obligations in the presence of this inference rule. But, given ideal world semantics, it does not follow from 'Op' that anyone has any obligation. Hence, the notion of satisfiability is not relevant. We have, then, a deontic logic which is a suitable specification for the logic of evaluation. What we do not have is an adequate means of formalizing statements which claim that a moral agent has a moral obligation which that moral agent might or might not satisfy. As a prolegomena to that, let us take a longer look at the various uses of 'ought'.

The use of 'ought' which so far seems clear is the use which I am calling the **optative** and, as we have seen, its function is evaluation. Some examples of this use are

(1) There ought to be no pain in the world,
(2) Al ought not have done that,
(3) You ought to be rewarded.

Not one of these examples identifies an agent and claims that the agent is bound by an obligation. The second example mentions an agent, but notice that it is in the past tense. It does not claim that Al *has* an obligation, but that he *had* one. The time for influencing Al's behavior is gone. Hence, (2) must be interpreted as an evaluation of an action which has already occurred. The third example mentions an agent, the person being addressed, but (3) is in the passive mood—the subject of the sentence is actually the object of the reward, not the agent doing the rewarding. The latter agent remains unspecified. Again, in the absence of any specified agent, it is difficult to construe (3) as a claim to the effect that there is an agent bound by some obligation. If it is not an evaluative use of 'ought', it is at least, clearly, an unclear case. There are plenty of clear cases on which to base a non-optative logic.

Non-optative senses do not seem to have any particular common bonds. We can dispose of one sense, however, that can be isolated quite easily from the others. This is the **predictive** sense. For example,

(4) This ought to make you feel better,
Replacing this switch ought to solve the problem.
The predictive sense is of no moral interest. Left over are sentences such as

(6) Phil is obligated to return it,
(7) You, Phil, ought to return it, right now!

It is not difficult to imagine (6) being said in a very impersonal, matter of fact way by someone far removed from the situation who has no interest in whether Phil returns it or not, but, still, is a person who is familiar enough with the situation to know that Phil is obligated to return it because he said he would (or something of that sort). On the other hand, (7) is the grammatical form in which a very similar thought might be expressed by someone who is very close to the situation and strongly emotionally tied to the outcome such as Phil's mother, wife, lover, close friend, or even the person who owns "it" and wants it back or desparetely needs it. There is certainly an active intention on the part of the speaker to influence Phil's behavior which is implied by (7) but not by (6). They both involve a positive evaluation but that is not their primary purpose. Their primary purpose is to express the fact that a moral agent is bound by an obligation. For this reason, they may not be treated straightforwardly as optatives because of Ross's paradox. Each one identifies an agent and implies that the agent is responsible for doing a certain action; that the agent is bound by an obligation to perform that action. They differ, perhaps, in the degree to which the speaker has a personal stake in what Phil does. But, then again, the difference may be a difference in kind. I am inclined toward
the latter view. (7) strikes me as being extremely close to, if not actually, telling Phil what to do. (6), however, has no such explicit "directive" element. (7) involves a greater force of some sort, probably, I think, imperative force. The similarities between (7) and Imperatives are quite striking. Notice that if the phrase 'ought to' is dropped from (7), the result is an ordinary imperative. Notice also that if statement (7) is changed from the second person that it loses that extra force. Imperatives, you will note, are always second person. (7) is the paradigm of what I have been calling prescriptives, and it is prescriptives that Castaneda tried, unsuccessfully, I believe, to explicate. I agree with him, however, that prescriptives involve an imperative element in some way. Just exactly what that way is will be exhibited shortly.

Both imperatives and prescriptives are typically second person and both are typically used to influence behavior. The distinction between prescriptive and optative uses of 'ought' has been exhaustively examined and defended by Paul Taylor. He says, "Prescribing is essentially and not merely incidentally a way of exerting an influence on someone's behavior. Expressing a value judgment, on the other hand, only incidentally has this function in most cases."1 'Ought' is normally used for prescribing and 'good' for evaluating, but 'ought' may also be used to evaluate. Prescriptive and evaluative sentences both have a speaker and an addressee (the person to whom the sentence is

1Taylor [1961].

2Taylor [1961], p. 226.
addressed) and prescriptive sentences prescribe only actions whereas evaluative sentences may evaluate anything including actions. There is also a special feature about the addressee of prescriptions. The addressee must be an agent who has the choice of doing or not doing the act being prescribed. Evaluations, on the other hand, may be addressed to anyone. Hence, 'John ought to repay his debt,' said to anyone other than John is not a prescription, it is an evaluation. This corresponds with Castaneda's concept of agency. But agency is not simply a matter of the prescription being second person and the agent being identical with the addressee. 'You ought not have done that,' is not a prescription since the time for choosing is gone. It is an evaluation of a past action of the addressee. Furthermore, 'Every student should turn work in on time,' is a prescriptive when the addressee is a student, otherwise it is evaluative. Similarly, 'We ought to do something,' is prescriptive when directed to the group referred to by 'we', but it is evaluative when spoken to someone outside the group. Notice that imperatives admit of these same deviations from the second person. For example, 'Everyone, sit down!' and 'Let's go!' In using 'ought' prescriptively we are advocates; in using it evaluatively, we are judges.

A prescription answers the question, 'What should I do?', and an evaluation answers the question, 'Why should I do it?'

There is, then, a clear cut distinction between the uses of 'ought' represented by (7), which I am calling prescriptives, and evaluative uses of 'ought' which I am referring to as optatives. But many of the characteristics of prescriptives do not seem to fit (6). It
clearly identifies an agent and claims that he is bound by an obligation, but it lacks the force of (7) which I have identified as imperative force. (7), as I have described it, implies (6) but the converse does not hold. (7) is a locution which can also express a prudential ought, an act, which if done, would benefit the doer of the act or, perhaps, the world. In this sense (7) can meaningfully be used in contexts in which no obligations are implied. It is not my purpose to discuss the prudential use, and I will not say a great deal about prescriptives either except to note that there is a sense of (7) in which it implies (6), a logical characteristic which is not terribly interesting. My main objective is to discuss locutions such as (6) and to describe their logical structure contrasting them with optatives.

To me, the most notable feature of (6) is that it has an aura of impartiality and objectivity. In fact, it looks very much like an ordinary factual claim. Of all the modern uses of 'ought', (6) seems to be the closest to the original use of the word\(^3\). Originally, the range of application of 'ought' was quite narrow, being restricted to duties arising out of contracts and agreements. The existence of an obligation presupposed the existence of a previous contract or agreement between two persons. Hence, the presence of an obligation was essentially a factual matter. Though an agreement is no longer necessary between two people to establish the existence of an obligation, (6) is certainly more like a factual claim than (7) is. To my mind, the main purpose of locutions such as (6) is to claim that an agent is in

\(^3\)My etymological information is taken from Brandt [1964].
Obligatives and prescriptives have one important logical feature in common. If $O(p)$ is the formalization of an obligative or a prescriptive, then it must not be possible to derive $O(p \lor q)$ from $O(p)$. Thus, there are three distinguishable senses of 'ought' which might be formalized in a complete deontic logic: the optative, the obligative, and the prescriptive sense. We have a candidate for the first, namely, ODL. An adequate logic for the last two must be able to deal with the notion of agency. This is the subject of the next section.

8.2 The Logical Requirements of Agency

Both prescriptives and obligatives require the mention of an agent. I am referring to this requirement as agency. My use of this term is now weaker than Castaneda's use of the term. He used it to refer to the agency of prescriptives (representing mandates) although he also seemed to be using agency to refer to the agent responsible for the action referred to in the statement involving agency. Contrary to Castaneda, I am claiming that both obligatives and prescriptives involve agency, and my use of 'prescriptive' is narrower than Castaneda's. He used the term to refer to directives, imperatives, orders, and entreaties. I am restricting it to uses of 'ought' that are second person (logically) and involve imperative force. Castaneda and I at least agree that what I call a prescriptive has something like an imperative
as an essential component. Our difference lies primarily in what we
take to be an adequate logic of prescriptives.

I propose, first of all, that imperatives be interpreted as
statements of the form

(8) You, a, see to it that p!

I believe that (8) adequately represents the features of imperatives in
general whether they are personified as commands, prescriptions, orders,
directives, or simply telling someone what to do. I believe that in all
cases, an imperative may at least be viewed as telling someone what to
do, and that we may also just as adequately interpret imperatives as
telling someone to see to it that some proposition becomes true or is
made true in virtue of an act performed by the agent. However, the log­
ic being developed is not a logic of change. It is assumed that all of
the wffs are atemporal, and if 'p' is false, then the order to see to it
that 'p' has been disobeyed. Even so, I believe that the resulting
action logic is sufficient at least for deontic purposes. If a person
has been told to see to it that 'p', and 'p' is false, then that person
has not done what they were told. The syntactic structure which I will
use for imperatives such as (8) is

(9) 1a[p],

where 'a' is the name of an agent, and 'p' is any wff of propositional
logic. The corresponding performance statement (retaining Castañeda's
term) is

(10) a[p]

which is read as
(11) \( \alpha \) sees to it that \( p \),
from which it may be inferred that 'p' is true.

Negation may be applied to (10) in two ways. First, the simple negative of (10) is

(12) \(-\alpha[p]\),
which is read 'It is false that \( \alpha \) sees to it that \( p \)' from which nothing follows about the truth or falsity of 'p'. The second way of applying negation to (10) is

(13) \( \alpha[-p] \),
which is read '\( \alpha \) sees to it that \( p \) is false' from which it follows that 'p' is false.

'\( \alpha[p] \)' and '\( \alpha[-p] \)' are viewed as descriptions of actions which require at least a minimal amount of attention on the part of the agent. Consequently,

(14) \( \alpha[p] \& \alpha[-p] \)
is logically false, but

(15) \( \alpha[p] \lor \alpha[-p] \)
is not logically true. (15) is false if \( \alpha \) is indifferent about whether 'p' is true or false. Looking at (14) and (15) suggests that before an imperative logic is built of expressions like '\( \alpha[p] \)', we must first work out a logic of "see-to-it-that". The logic which I will present as a candidate for the logic of see-to-it-that was developed by von Wright who calls it an action logic. I shall refer to it as the logic of conscious action (CL) for reasons explained in the next section.
8.3 The Logic of Conscious Actions

Von Wright presented his action logic in a paper read before the Creighton Club. He then constructed a deontic logic on top of his action logic as part of the same paper. The resulting deontic logic is not as rich as Hintikka's deontic logic, and it is seriously hampered by von Wright's restriction of the variables to verb phrases, and by his terribly cumbersome semantics. In fact, all of the arguments presented against \( P \) in Chapter III apply to this new deontic logic, as well as Hintikka's arguments about the need for quantifiers. The virtue of his new deontic logic, and this single virtue overshadows many of the faults, is that it contains neither Ross's paradox nor the other paradoxes based on equivalent expressions. This happy circumstance is not due to features of the deontic logic but on the interesting nature of the underlying action logic. I have taken the liberty of using the formal characteristics of von Wright's action logic for my own purposes. Expunged of von Wright's uneasiness over propositions and clothed in a more streamlined syntax, it strikes me as being an admirable basis for an imperative logic. To my knowledge, von Wright has not put it to this use. I have also abandoned his truth tables in favor of a tree procedure which, by now, the reader should find quite familiar.

The syntactic elements of CL are the standard elements of propositional logic plus lower case Greek letters which function as the operator

\[
\text{(16) The agent, } \alpha, \text{ sees to it that...}
\]

\(^4\) von Wright [1973].
A string of elements is well-formed if and only if it is formed in accordance with one of the following two rules:

(FR1) If \( p \) is a wff of propositional logic and \( \alpha \) is an agent operator, then \( \alpha[p] \) is a wff of CL.

(FR2) Truth functions of wffs of CL are wffs of CL.

In any expression of the form \( \alpha[p] \), \( '[p]' \) is the purview of \( \alpha \) and provably equivalent expressions may not be substituted for one another within the purview of an agent. Replacement may take place only via definitions:

(Def. [v]) \( \alpha[p \lor q] = df \alpha[-(-p \land -q)] \),

(Def. [\( \lor \)]) \( \alpha[p \lor q] = df \alpha[-p \lor q] \),

(Def. [\( \equiv \)]) \( \alpha[p \equiv q] = df \alpha[(p \land q) \lor (-p \land -q)] \).

If we allow the substitution of provable equivalents, then \( '[p]' \) may be replaced with its conjunctive normal form, \( '(p \lor q) \land (p \lor -q)' \). Hence, \( '[\alpha[p]]' \) implies \( '[\alpha[(p \lor q) \land (p \lor -q)]]' \) and the latter implies \( '[\alpha[p \lor q]] \land [p \lor -q]' \) which, in turn, implies \( '[\alpha[p \lor q]]' \). Consequently, \( '[\alpha[p \lor q]]' \) may be derived from \( '[\alpha[p]]' \) and CL is not free of Ross's paradox.

The distinctive feature of CL is the manner in which the agent operator distributes over truth functions. First, some terminology. If a modal operator distributes over a connective in the manner in which \( 'Q' \) distributes over \( '+' \) in

\[
(17) \quad Q(p + q) \leftrightarrow (Qp + Qq),
\]

I shall say that \( 'Q' \) distributes over \( '+' \) directly. If, on the other hand, the operator \( 'Q' \) distributes over the disjunctive normal form of
'p + q', I shall say that 'Q' distributes over '+' indirectly. In CL the agent operator distributes over '&' directly and it distributes over 'v' indirectly. In other words, CL has the following distribution laws:

(DL1) \( a[p \lor q] = (a[p \land q] \lor a[p \land \neg q] \lor a[\neg p \land q]) \),

(DL2) \( a[p \land q] \rightarrow (a[p] \land a[q]) \).

The indirect distribution of 'a' over 'v' requires a semi-intentional reading of 'a[p]', otherwise a fully intentional reading leads to paradoxes. What I mean by this is that we may not interpret 'a[p]' as 'a intends that p be true.' But we may not treat it as the completely non-intentional 'a does p' either. Interestingly enough, 'a[p]' is also semi-intensional (with an 's' instead of a 't') because, as we shall see, the truth-value of 'a[p]' is not entirely determined by the truth value of 'p'. If 'p' is false, then of course a does not see to it that 'p'. But if 'p' is true, then the truth value of 'a[p]' is undetermined. Furthermore, as noted above, CL does not allow the substitution of logically equivalent expressions within the purview of an agent.

I will now show why a fully intentional (with a 't') interpretation of 'a[p]' is not possible. In virtue of the rules of CL,

(18) \[ a[p \lor q] \]

\[ \neg a[p] \]

Hence: \( a[q] \)

is a valid inference. At the beginning of a basketball game the referee intends that either one team or the other team gains possession of the
ball at the jump, i.e.,

(19) \[\alpha[p \lor q],\]

the first premise of the argument to be constructed, is true. Suppose further that the referee is impartial. It follows that it is false that he intended team 'p', let us say, to get the ball, i.e.

(20) \[-\alpha[p],\]

the second premise, is true. Then, according to CL, the referee intended that team 'q' get the ball at the jump. But on an intentional interpretation this should not follow. On an intentional interpretation an impartial referee should be able to intend that one team or the other gains possession of the ball but neither intends that one team nor the other team gets the ball. So, CL will not do for explicitly intentional acts.

But a completely non-intentional interpretation such as 'α does p' for 'α[p]' requires that 'α' distribute directly over 'v'. We must suppose that α is at least conscious of the fact that his action has 'p' as a result; he must be aware of what he is doing and that it will lead to 'p'. It is at least a conscious action. This, I think, is captured quite nicely by the reading

(21) \[\alpha \text{ sees to it that } p,\]

which is hereby adopted as the official reading of 'α[p]' .

The interpretation of

(22) \[\alpha[p \land q]\]

is straightforward. If α sees to it that 'p' and 'q', then α sees to it that 'p' and α sees to it that 'q'. The converse does not hold.
If 'α[p]' and 'α[q]' are true it does not follow that 'α[p ∧ q]' is true since there need not be any conscious or deliberate connection between α seeing to it that 'p' and α seeing to it that 'q'. There is one last rule of CL. 'Sees to it that...' is what philosophers in recent years have come to call an achievement verb. From the fact that α sees to it that 'p', it follows that 'p'. Hence, 

(R1) \[ α[p] \rightarrow p \]

is a theorem of CL. (R1) was not a rule in von Wright's original act logic because he interpreted the variables as verb phrases. This means that von Wright reads 'p' as 'is going to the store'. Hence 'α[p]' is read as 'α is going to the store.' This reading does not seem compatible with the indirect distribution of the agent operator over 'v', because von Wright's reading is non-intentional. Despite the fact that von Wright's interpretation does not coincide well with his syntactic rules, his motive for distributing 'α' over 'v' indirectly is quite clear—it avoids Ross's paradox whereas distributing directly does not. To be precise, 

(23) \[ α[p] \rightarrow α[p ∨ q] \]

is not a theorem of CL. Having that formal result is one thing; providing an adequate interpretation is another. In that respect, I believe that CL represents a significant improvement over von Wright's original act logic.

8.4 Tree Rules for CL

As before, the negation of a CL-valid formula does not have a consistent model. CL-models satisfy all of the conditions of proposi-
tional logic plus three new conditions:

(C.A)  If $a[p]$ is true, then $p$ is true.

(C.[&])  If $a[p \& q]$ is true, then $a[p]$ is true and $a[q]$ is true.

(C.[v])  $a[p \lor q]$ is true if and only if either $a[p \& q]$ is true,

or $a[p \& \neg q]$ is true, or $a[\neg p \& q]$ is true.

Truth trees representing CL-models may be constructed by adding the
rules listed in Table 7 to the tree rules for propositional logic.

As an illustration we shall see how the rules show that

(26)  $a[p] \rightarrow a[p \lor q]$

is not a theorem of CL. See Figure 12. The tree indicates that the
counter-example to (26) is a case in which $a$ sees to it that '$p'$ is
true, but is indifferent about the occurrence of '$p'$ together with
either '$q'$ or '$\neg q'$. In other words, the truth or falsity of '$q'$ is
not part of the aim of $a$'s deliberate action in this case. I shall
take this occasion to list a few theorems of CL and then turn to the
development of an imperative logic.

Distribution Laws:

(DL3)  $a[\neg(p \& q)] \equiv (a[\neg p \& -q] \lor a[\neg p \& q] \lor a[-p \& q])$

(DL4)  $a[-(p \lor q)] + (a[-p] \& a[-q])$

(DL5)  $-a[-(p \& q)] \equiv (-a[p \& -q] \& -a[\neg p \& q] \& -a[\neg p \& -q])$

(DL6)  $-a[p \lor q] \equiv (-a[p \& q] \& -a[p \& -q] \& -a[\neg p \& q])$

Entailments:

(T1)  $(a[p] + a[q]) + (p + a[q])$

(T2)  $a[p \rightarrow q] + (a[p] + a[q])$

(T3)  $a[p] + (p \lor q)$
\[ \sqrt{ \alpha[p \lor q] } \]

\[
\begin{array}{c}
\alpha[p] \\
\alpha[p] \\
\alpha[-p] \\
\alpha[q] \\
\alpha[-q] \\
\alpha[q] \\
\end{array}
\]

\[ \sqrt{ \alpha[-(p \land q)] } \]

\[
\begin{array}{c}
\alpha[p] \\
\alpha[-p] \\
\alpha[-p] \\
\alpha[-q] \\
\alpha[q] \\
\alpha[-q] \\
\end{array}
\]

\[ \sqrt{ \alpha[p \land q] } \]

\[
\begin{array}{c}
\alpha[p] \\
\alpha[-p] \\
\alpha[-q] \\
\end{array}
\]

\[ \sqrt{ \alpha[-(p \lor q)] } \]

\[
\begin{array}{c}
\alpha[p] \\
\alpha[-p] \\
\alpha[-q] \\
\end{array}
\]

\[ \sqrt{ \alpha[p] } \]

\[
\begin{array}{c}
\alpha[p] \\
\alpha[-p] \\
\alpha[-q] \\
\end{array}
\]

**TABLE 7: TREE RULES FOR CL**

<table>
<thead>
<tr>
<th>Rule</th>
<th>Description</th>
<th>Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \neg (\alpha[p] \rightarrow \alpha[p \lor q]) )</td>
<td>Assume</td>
</tr>
<tr>
<td>2</td>
<td>( \alpha[p] )</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>( \alpha[p \lor q] )</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>( \alpha[p \land q] )</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>( \alpha[p \land \neg q] )</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>( \alpha[\neg p \land q] )</td>
<td>3</td>
</tr>
</tbody>
</table>

**FIGURE 12**
As specified earlier, imperatives such as

(25) You, α, see to it that p!

will be formalized as

(26) !α[p].

The difference between (26) and

(27) !α[-p]

is the difference between commanding and countermanding mentioned earlier. The difference between (26) and

(28) -!α[p]

is the difference between commanding and simply stating that a particular command, for one reason or another, is not in force. Hence, IL is capable of formalizing the two types of imperative negation that, it was noted, Castaneda could not formalize.

Accordingly, IL is an extension of CL which means that all rules and wffs of CL are rules and wffs of IL. To produce IL, two additional formation rules are necessary:

(FR3) If p is a wff of propositional logic and α is an agent operator, then !α[p] is a wff of IL.

(FR4) Truth functions of wffs of IL are wffs of IL.
To the tree rules for CL we add rules which are derived from the following conditions for '1':

(C.-l) If \(-\alpha[p]\) is true in some world, \(W_i\), then there is an imperative world alternate to \(W_i\) in which \(-\alpha[p]\) is true.

(C.l) If \(\alpha[p]\) is true in some world, \(W_i\), then \(\alpha[p]\) is true in all imperative worlds alternate to \(W_i\) and there is at least one world alternate to \(W_i\).

Suppose that the following order is in force:

(29) \(\neg\alpha[p \& q]\),

and suppose that \(\alpha\) sees to it that 'q' but fails to see to it that 'p' because, \(\alpha\) argues, he is under no order to see to it that p:

(30) \(\neg\alpha[p]\).

I believe that such reasoning is properly resisted. (C.-l) allows us to show that (29) and (30) are inconsistent. One might reject (C.-l) in favor of

(31) \(\alpha[p \& q] \to (\alpha[p] \& \alpha[q]).\)

However, condition (C.-l) will show that (31) is an imperative tautology. I have no idea what sort of semantic condition would be necessary to preserve (31) in the absence of (C.-l), nor do I see any reason for not adopting (C.-l) if one is going to go so far as to adopt (31).

I can think of no counter-intuitive valid formulas produced by (C.-l).

An argument against (31) could claim that '\(\neg\alpha[p] \& \alpha[q]\)' does not follow from '\(\neg\alpha[p \& q]\)' since 'p' and 'q' may be connected in such a way that both must be done together. Hence, the individual commands, '\(\neg\alpha[p]\)' and '\(\alpha[q]\)' simply do not make sense, e.g., an order to clean up
and be quick about it implies neither an order to clean up (at an unspecified pace) nor an order to be quick about it. The attentive reader may notice a certain *déjà vu* character to this argument.

Similar arguments were discussed in connection with utilitarianism (section 2.6) and interpreting variables as act-types (section 6.1). I have argued that 'p' and 'q' must be interpreted as being independent actions (or descriptions of actions or results of actions). Hence, '!(α[p & q])' does not adequately represent all that is said by (32).

A full treatment requires a quantificational treatment in which (32) is understood as ordering α to see to it that there is a single performance which is an instance of cleaning up and which is also an instance of being quick.

In an imperative such as '!(α[p])' I shall call 'α[p]' the direct intention or direct object of the imperative, and 'p' by itself is the indirect intention or indirect object of the imperative. If the direct object of an imperative is true, then the imperative is obeyed or satisfied. The indirect object sometimes satisfies the imperative, but not always. If we tell someone to shut the door, we may accept someone else shutting the door. In the armed services, however, the indirect object of a command is virtually never sufficient to satisfy a command. Later, this distinction will be applied to deontic formulas, e.g., '0α[p]'.

The word 'Intention' is used because '!(α[p])' is intentional in the same way in which it was argued that 'α[p]' is not intentional. Like the basketball referee, one may intend that '(p v q)' yet neither intend 'p' nor intend that 'q'. Similarly, it may be imperative
(ordered or directed) that 'a[p v q]' yet it is neither imperative
(ordered or directed) that 'p' nor imperative (ordered or directed)
that 'q'. To put it another way,

(33)  \[ \neg \alpha[p v q] \]
     \[ \neg \alpha[p] \]
     Hence:  \[ \neg \alpha[p] \]
is not a valid inference. This is easily seen when \[ \neg \alpha[p] \]
is interpreted propositionally. For example:

(34)  \[ \alpha \text{ is ordered to fix it or replace it,} \]
     \[ \alpha \text{ is nor ordered to fix it,} \]
     \[ \alpha \text{ is not ordered to replace it,} \]
is a consistent set of statements. Despite the fact that \( \alpha \) has been ordered to fix it or replace it and that no order has been issued to
fix it, it does not follow that an order has been issued to replace it.
The forest illustrated by Figure 13 shows that (27) is invalid and
serves as an illustration of how the rules are applied to trees.

IL can show that

(35)  \[ \text{You, Al, fix it or replace it!} \]
     Hence:  \[ \text{You, Al, replace it if you don't fix it!} \]
is valid and that

(36)  \[ \text{You, Al, fix it! or You, Bob, replace it!} \]
     Hence:  \[ \text{You, Bob, replace it if it isn't fixed!} \]
is invalid. On Castaneda's imperative logic both are invalid. Further-
more, IL can be used to show what further assumptions are necessary in
order to show that (36) is valid. The test sentence for (35) is:
\[ W_0 \]

1. $\sqrt{1 \alpha [p \lor q]} \ast \text{Assume}$
2. $\sqrt{-\alpha [p]} \text{ Assume}$
3. $\sqrt{-\alpha [q]} \text{ Assume}$

\[ W_1 \]

1. $-\alpha [p] \quad W_0 - 2$
2. $\sqrt{\alpha [p \lor q]} \quad W_0 - 1$
3. $\alpha [p] \quad \alpha [p] \quad \sqrt{\alpha [-p]} \quad 2$
4. $\alpha [q] \quad \alpha [-q] \quad \sqrt{\alpha [q]} \quad 2$
5. $X \quad X \quad -p \quad 3$
6. $\quad \quad \quad q \quad 4$

\[ W_2 \]

1. $-\alpha [q] \quad W_0 - 3$
2. $\sqrt{\alpha [p \lor q]} \quad W_0 - 1$
3. $\alpha [p] \quad \sqrt{\alpha [p]} \quad \alpha [-p] \quad 2$
4. $\alpha [q] \quad \sqrt{\alpha [-q]} \quad \alpha [p] \quad 2$
5. $X \quad p \quad X \quad 3$
6. $\quad \quad \quad -q \quad 4$

\text{FIGURE 13}
(37) \( !a[p \lor q] + !a[-p \lor q] \),
and the test sentence for (36) is
(38) \( (1a[p] \lor !B[q]) \lor !B[-p \lor q] \),
using the following logical vocabulary: 'a' = 'Al'; 'B' = 'Bob'; 'p' = 'Al fixes it'; 'q' = 'Bob replaces it'. I will omit the trees. They show that (37) is logically true, but (38) is not. Castaneda must formalize (35) and (36) with a single expression, viz.,
(39) \( (A \lor B) \rightarrow (-p \lor B) \),
which, according to his rules, is not valid. The reason that (38) fails to be valid is the simple fact that no single agent (especially Al) is given complete responsibility for seeing to it that Al fixes it or Bob replaces it. If the premise of (36) is interpreted such that Al and Bob have the joint responsibility for seeing to it that Al fixes it or Bob replaces it, then the formal representation would be:
(40) \( (1a[p \lor q] \& !B[p \lor q]) + !B[-p \lor q] \),
which is a theorem of IL. Needless to say, Castaneda's logic does not contain the equipment to express the relationship exhibited by (40).

How should expressions such as
(41) \( 1a[p] \lor !B[q] \),
which express the content of Castaneda's 'A v B', be interpreted?
A clue may be taken from the connective. It is a truth function, thus, the clearest reading will be a truth functional one, viz.,
(42) Either it is imperative that a see to it that p or it is imperative that B see to it that q.
On this reading, (41) is not an imperative. Hence, I have aligned my-
self with those who believe that the logical relations of imperatives must be analyzed in terms of improper imperatives. (See section 1.2.) Castaneda's logic cannot capture the difference between arguments (35) and (36) because it cannot formalize the difference between (41) and (43)

$$\text{lα}[p \lor q].$$

IL, however, is sensitive enough to detect the difference and to explain the precise logical requirements for there to be a valid interpretation of (36).

Among the logical requirements of commands are:

$$\text{(44)} \quad \text{lα}[p \land q] \equiv (\text{lα}[p] \land \text{lα}[q]),$$

and

$$\text{(45)} \quad -(\text{lα}[p] \land \text{lα}[-p]).$$

But we do not find

$$\text{(46)} \quad \text{lα}[p \lor q] \equiv (\text{lα}[p \land q] \lor \text{lα}[p \land -q] \lor \text{lα}[-p \land q])$$

among the theorems of IL which is just as well. If we did, then (33) would be a valid inference.

8.6 Quantifiers

It is common for action logics to distinguish between doing and refraining or commission and omission. Von Wright, for example, has introduced both types of operators\(^5\). CL draws this distinction with quantifiers. This requires a slight modification of the formation rules to allow predicate functions and quantifiers within the scope of agent operators. Following Hintikka, the individual variables range over act

\(^5\)Von Wright [1963], p. 56.
individuals. Quantifying over acts allows us to utilize the results of the discussion of section 6.2.

Two reasonable candidates for formalizing acts of commission are

(47) \((\exists x)\alpha[Ax]\)

and

(48) \(\alpha[(\exists x)Ax]\).

They are not equivalent. Suppose that the universe of discourse is restricted to the two acts 'a' and 'b'. Then (47) is equivalent to

(49) \(\alpha[Aa] \vee \alpha[Ab]\),

and (48) is equivalent to

(50) \(\alpha[Aa \vee Ab]\)

which is equivalent to

(51) \(\alpha[Aa \& Ab] \vee \alpha[Aa \& -Ab] \vee \alpha[-Aa \& Ab]\).

You can easily see that the expansion of (48) becomes quite large as the number of individual actions in the universe of discourse grows. For \(n\) acts, the number of disjuncts in the expansion will equal \(2^n - 1\).

Compare (51) with the disjunctive Boolean expansion of (49):

(52) \((\alpha[Aa] \& \alpha[Ab]) \vee (\alpha[Aa] \& -\alpha[Ab]) \vee (-\alpha[Aa] \& \alpha[Ab])\).

Notice that if '\(\alpha[Ab]\)' is true and '\(\alpha[-Aa]\)' and '-\(\alpha[Aa]\)' are both false, then (51) is false but (52) is true. In order for (51) to be true, \(\alpha\) must see to more than just 'Aa', he must also consciously see to it that 'Ab' is either true or false rather than letting the chips fall where they may. Hence, (48) implies a conscious action of a wider scope than (47) does. Accordingly, we must provide semantic rules which will
render (47) a logical consequence of (48) but which avoid the converse. And if \( 'a[Aa]' \) is true, then we may infer (47) but not (48).

The difference between (47) and (53)

\[(x)\alpha[Ax]\]

is straightforward. The difference between (48) and (54)

\[\alpha[(x)Ax]\]

is almost as straightforward. Since the expansion of (54) in a domain of two actions is (55)

\[\alpha[Aa \& Ab],\]

it should be evident that (55) implies (51), hence (54) implies (48). The converse does not hold.

We may also note that (54) implies (53). Again, the converse does not hold. The expansion of (53) in a domain of two actions is (56)

\[\alpha[Aa] \& \alpha[Ab].\]

According to the rules for CL, (55) implies (56), hence (54) implies (53). This difference can be accounted for by noting that (53), unlike (54), does not imply any sort of "conspiracy" or "plot" on the part of \( \alpha \) to see to it that every action is an instance of 'A'. It has just happened that for every action which has come along, \( \alpha \) has seen to it that it is an instance of 'A'. This description would also follow from the execution of a plot by \( \alpha \) to see to it that every action is an instance of 'A'. Hence, (54) implies (53).

In order to show that

\[(57) \quad \alpha[(x)Ax] + (x)\alpha[Ax],\]

\[(58) \quad \alpha[(x)Ax] \rightarrow \alpha[(\exists x)Ax],\]
(59) \[ \alpha[Aa] \to (\exists x)\alpha[Ax], \]

and

(60) \[ \alpha[(\exists x)Ax] \to (\exists x)\alpha[Ax] \]

are theorems of CL, we must introduce two new semantic conditions for quantifiers which fall within the purview of an agent. These are:

(C.[E]) If \( V(\alpha[(\exists x)\phi x], W_i) = T \), then \( V(\alpha[\phi \mu], W_i) = T \) for some member \( \mu \) of \( W_i \), otherwise \( V(\alpha[(\exists x)\phi x], W_i) = F \).

(C.[U]) If \( V(\alpha[(x)\phi x], W_i) = T \), then \( V(\alpha[\phi \mu], W_i) = T \) for every member \( \mu \) of \( W_i \), otherwise \( V(\alpha[(x)\phi x], W_i) = F \).

And for purposes of making changes within the purview of an agent we suppose that

(Def.Q) \[ \alpha[(x)\phi x] =_{df} \alpha[-(\exists x)-\phi x]. \]

Please note that these new semantic conditions do not allow the removal of a quantifier if the quantifier falls within the scope of any logical constant other than an agent term.

CL is capable of expressing both weak commission and strong with (47) and (48), respectively. The difference between them is something like the difference between

(61) There is at least one instance of Adam seeing to it that the bell is rung,

and

(62) Adam sees to it that the bell is rung at least once.

In other words, in the latter, the number of times that the bell is rung is part of what \( a \) has in mind in doing the action. They are similar in that they both imply that \( a \) sees to it that the bell is rung.
The difference becomes clearer when we compare the translations of

(63) There was just one act which was an act of Adam
deliberately ringing the bell.

and

(64) Adam deliberately rang the bell just once.

The translations should be

(65) \( (\exists x)(\alpha\neg Ax \land (y)(\alpha\neg Ay \land (x = y))) \)

and

(66) \( \alpha[\neg (\exists x)Ax \land (y)(Ay \land (x = y))] \),

respectively.

Likewise, weak and strong omissions are formalized by

(67) \( (x)\neg Ax \)

and

(68) \( \neg (x)Ax \).

Again, the difference is that (67) is compatible with things just having worked out that \( \alpha \) decided in each case to make that case an instance of 'A', whereas (68) requires that every action fail to be an instance of 'A' as the result of a conscious effort by \( \alpha \) to see to it that every action fails to be an instance of 'A'.

(67) and (68) represent a sense of omission which is somewhat stronger than the usual sense of omission. This is because 'Aa' need not be an action done by \( \alpha \). We decide whether it is or not by inspecting the predicate. A permissible reading of 'Ax' is 'x is an instance of stealing' as well as 'x is an instance of \( \alpha \) stealing'. Given the first reading, (68) says that \( \alpha \) sees to it that no one steals; that
there is no action performed by anyone which is an instance of stealing. This is not the usual sense of omission since the action which $a$ does to see to it that no one steals may be something like declaring martial law, killing all thieves, or confiscating or destroying all property. But this is not particularly paradoxical, and any restriction to the effect that the actions which fall within the purview of an agent must be actions of that agent leads to any unnecessary decrease in the power of CL.

Applying '!' to the two forms of omission gives

$(69) \quad l(x)a[\neg A\chi]$

and

$(70) \quad la[(x)-A\chi].$

I do not think that $(69)$ should be well-formed for the following reasons. First, I have the feeling that $(69)$ does not address itself to the agent, but to the totality of actions. It is clear in $(70)$ that $a$ is ordered to see to it that '(x)-A\chi'. But it is not clear that $a$ is the commanded agent in $(69)$. Secondly, $(69)$ and $(70)$ do not appear to capture any important difference. So I have simply stipulated that $(69)$ is ill-formed. In general, it seems that no syntactic element should come between '!' and 'a'.

Is it possible to move the quantifier in $(70)$ to the left of '!'? I don't think so. When one gives a command, its direct intention is that there be an action of a certain sort, not that an existing action is to have a certain characteristic added to it, which, I think, is impossible. It is absurd to say 'Do it carefully!' and, when asked
what 'it' refers to, to say: 'Writing the paper you finished last week.' For a command to be "felicitous" it must be about actions which do not exist, otherwise it is most difficult to suppose that commands and imperatives are recommendations that there be actions of various sorts, and it is difficult to see how they can be interpreted any other way.

These are the considerations which have led to the constraints placed upon the use of act quantifiers which are built into the next formation rule:

\[(FR5) \quad \text{If } A \text{ is a wff which contains no occurrences of '!}',\]
\[\text{then } (x)A \text{ is a wff.}\]

Except for one further complication, the formal machinery required for an improved deontic logic (IDL) is complete. That further complication is quantification over agent individuals. It does not affect, in any significant way, the relationships between obligation, permission, and prohibition, the next topic to be discussed.
CHAPTER IX.
AN IMPROVED DEONTIC LOGIC

9.1 The Syntax of IDL

IDL is simply the result of combining the quantified action logic and imperative logic presented in the previous chapter with the optative deontic logic defended in Chapter VI. To be complete, I shall add one formation rule to those listed in the previous chapter:

(FR6) If A is a wff, then OA is a wff.

As always, the negation of an IDL-valid formula has no consistent model. IDL models satisfy all of the conditions listed in the previous chapter plus (C.P), (C.O), and (C.Ref), plus one new condition which will be described shortly.

My intention is to formalize the senses of 'ought' distinguished in section 8.1 as follows:

Optative: Op,
Obligative: Oα[p],
Prescriptive: O1α[p].

Clear cut cases of the optative are those in which 'p' contains no agent terms. But, strictly speaking, any expression of the form '0p' is an optative use on the semantics of IDL whether 'p' contains an agent term or not. The operator '0' represents the basic deontic notion and it plays a role in the formalization of all deontic uses of 'ought'. The
difference between obligatives and prescriptives described in section 8.1 is captured quite nicely by their formalizations. The imperative force of prescriptives is built into their formal representation by construing them as positive evaluations of imperatives. On the semantics of IDL, however,

\[ (1) \quad 0!\alpha[p] \rightarrow !\alpha[p] \]

is not logically true. So I am not sure that prescriptives tell a person what to do (as '!\alpha[p]' clearly does). They surely tell a person what he should do. If you choose to say

\[ (2) \quad \text{You, Adam, ought to go to the store,} \]

instead of

\[ (3) \quad \text{You, Adam, go to the store!} \]
you have weakened the imperative. Unlike saying (3), you open yourself up to the possibility of an argument when you say (2). An appropriate response to (2) is 'You're wrong', but as a response to (3), 'You're wrong' makes no sense. If Adam fails to go to the store in the face of (3), Adam has been disobedient. If Adam fails to go to the store when (2) has been said to him, he is exercising what remains his prerogative.

In section 8.1, I said that a prescriptive implies its corresponding obligative, i.e.,

\[ (4) \quad 0!\alpha[p] \rightarrow 0\alpha[p] \]

is logically true. This necessitates a further semantic assumption, \textit{viz.}, All imperatives are obeyed in ideal worlds. If an imperative is in force in an ideal world, then that imperative has been positively valued. If a command is positively evaluated, it seems that we are also
positively evaluating the direct object of the command. This is the content of

(5) \( O(!\alpha[p] \rightarrow \alpha[p]) \).\(^1\)

This principle, the principle that orders ought to be obeyed, is a deontic consequence—not a logical one. Hence IDL requires an additional semantic condition:

\[ (C.01) \quad \text{If } !A \text{ is true in an ideal world, then } A \text{ is also true in that world.} \]

Other expressions which owe their IDL-validity to this condition are

(6) \( O(!\alpha[p] \rightarrow p) \),

(7) \( O!\alpha[p] \rightarrow Op \).

Lastly, the translations satisfy another requirement established in section 8.1, and that is that the formalizations of prescriptives and obligatives must avoid Ross's paradox. It is not difficult to establish that

(8) \( 0\alpha[p] \rightarrow 0\alpha[p \lor q] \)

and

(9) \( 0!\alpha[p] \rightarrow 0!\alpha[p \lor q] \)

are not logically true. It might be replied that

(10) \( 0\alpha[p] \rightarrow O(\alpha[p] \lor \alpha[q]) \)

\(^1\)Another reasonable semantic notion is that the object of a positively evaluated command is merely permitted. In this way \( O!\alpha[p] \) can be interpreted as a prudential 'ought' advocating that \( \alpha \) see to it that \( p \) even though \( p \) is not obligatory, but it is permitted and seeing to it would lead to personal gain on the part of \( \alpha \), or simply make the world a better place to live. See section 8.1.
and

\[(11) \quad 0!α[p] \to 0(0α[p] \lor 0α[q])\]

are theorems of IDL (which they are), hence, Ross's paradox is not avoided and IDL is no better than its predecessors. Fortunately, this move may be resisted. The consequent of (10) is not an obligative. The formal requirement of an obligative is that it be of the form '0α[p]' . The deontic operator must apply directly to the agent operator. This is a stricter requirement than simply asking that an agent term fall within the scope of the deontic operator. The same comment applies, *mutatis mutandis*, to (11). It is not a prescriptive since the main logical operator of the expression to which the deontic operator is applied is not an agent term. The consequents of (10) and (11) must be interpreted as optatives. It was also noted in section 8.1 that prescriptives and obligatives both state the existence of obligations and that these obligations must be expressed in terms of the agent bound by them. These are the kinds of obligations that we may sensibly speak of satisfying. The notion of satisfiability is at the heart of Ross's paradox and other classical difficulties in deontic logic. IDL can handle this notion quite easily, and, as far as I can tell, without paradoxical consequences. A prescriptive, '0lα[p]', or an obligative, '0α[p]', is strongly satisfied if 'α[p]' is true and weakly satisfied if 'p' is true. Since the arguments of (10) and (11) are neither prescriptives nor obligatives, they are not the sort of thing that can be satisfied, so there is no way to get Ross's paradox started.
IDL is also free of the derived obligation paradox, i.e., neither

(12) $0\alpha[p] \rightarrow 0\alpha[-p \rightarrow q]$

nor

(13) $0!\alpha[p] \rightarrow 0!\alpha[-p \rightarrow q]$

are theorems of IDL. This has the happy consequence of allowing the expression

(14) $0\alpha[p \rightarrow q]$

to express a *prima facie* obligation since '0$\alpha[r \rightarrow q]' does not follow from '0$\alpha[-r]'. IDL thus avoids the difficulty of HDL noted in section 6.3 that trivial *prima facie* obligations cannot be distinguished from important ones. In light of other comments in section 6.3, an actual obligation will be formalized as '0$\alpha[p]''. There is no harm, however, in referring to 'p + 0$\alpha[p]' as an actual obligation since an actual obligation obtains when 'p' is true.

9.2 The Uses of 'Ought' Formalized

In Chapter VI there appears a chart detailing the logical relationships of the various combinations of deontic operators and quantifiers. On the basis of the logical relationships exhibited there, I rejected Hintikka's view that the relations between the three moral categories under discussion were more complex than previously imagined. That claim should be understood as referring to the deontic categories formalizable in HDL. By now it should be clear that IDL is capable of
formalizing much more.

We may now distinguish between positive duties and negative duties. A positive duty is an obligation to commit some action (e.g., 'You ought to push this button at 3 p.m.') and a negative duty is an obligation to refrain from a certain type of action (e.g., 'You ought never lie.'). The positive prescriptive use of 'ought' is captured by

\[ 0!a[(\exists x)Ax] \]

which implies the corresponding positive obligative:

\[ 0a[(\exists x)Ax] \]

which, in turn, implies

\[ 0(\exists x)a[Ax] \]

which is neither a prescriptive nor an obligative. I shall refer to (17) as a performance evaluative use of 'ought'. The deontic operator does not apply directly to the agent but to the quantifier, so it should be read as 'It ought to be the case that there is an action of a certain sort which is deliberately done by \( a \).' (17) positively evaluates there being such an act. (16), on the other hand, positively evaluates \( a \)'s entire plan of deliberate action (seeing, for example, that these three actions (and no others) are instances of 'A'). Lastly, (17) implies

\[ 0(\exists x)Ax \]

which simply evaluates positively the existence of an act of a certain sort. This is the act evaluative use of 'ought'.

In addition to these four forms, there are two others which are logically weaker. These are

\[ (\exists x)0a[Ax] \]
The difference between (17) and (19) and their similar counter-parts, (18) and (20), is the distinction drawn by Paul Taylor between post eventum and ante eventum uses of 'ought'. The difference depends on whether the action being evaluated occurs before or after the act of evaluation. Those uses of 'ought' which evaluate actions which occur prior to the act of evaluation are post eventum uses. Those uses which evaluate actions occurring after the evaluation are ante eventum. In addition, (18) and (20) differ with respect to what they claim exists in the world. The former does not claim that an act exists in the world. The latter does assert that there is at least one action in this world, and on my view an existing act is a fait accompli. Hence, the evaluation made by (20) is a post eventum evaluation; specifically, a post eventum evaluation of an action. (19) is a post eventum evaluation of a performance. Likewise, (18) is an ante eventum evaluation of an action, and (17) is an ante eventum evaluation of a performance.

It should be evident that I am understanding performances as being deliberate actions. This does not strike me as being unreasonable. I believe that this is the basis of the difference between human performances and mechanical events. A human performance is at least a conscious action; one which the agent is consciously aware of, and it can range all the way from the exercise of a habit, which the agent, however dimly, consciously allows his body to do, to a fully conscious,

\[^2\]Taylor [1961], p. 195.
deliberate attempt to mold the environment to suit the purposes and desires of the agent. And if the general thesis that I am defending is correct, there is a fundamental logical difference between human performances and mechanical events.

Let us now turn to negative duties. First, a negative prescriptive is

\[ (21) \quad 0\alpha[(x)-Ax] \]

which implies the corresponding obligative:

\[ (22) \quad 0\alpha[(x)-Ax], \]

a statement to the effect that one has a negative duty toward a certain kind of action. The obligative implies

\[ (23) \quad 0(x)\alpha[-Ax], \]

which formalizes an evaluative use of 'ought'. Unlike the corresponding formalizations of positive duties which employ the existential quantifier, we may not distinguish between (23) and

\[ (24) \quad (x)0\alpha[-Ax] \]

on the basis of the ante eventum-post eventum distinction because (23) logically implies (24). Clearly, then (23) must be interpreted as an evaluation of all actions including those which occur both before and after the act of evaluation. (24) is at least an evaluation of all actions which have already occurred. We can only hope that it will continue to be true of new actions as they are added to the world. If falsifying cases pop up, then we must discard (24) in favor of other deontic truths. To fully appreciate the difference between (23) and (24) we must appeal to the rules for 'O' and 'P'. On this basis, we may see that (24) is telling us what is true of all (past) actions in this world which,
incidentally, include such things as torture, cheating, infanticide, and just plain nastiness. On the other hand, (23) is not about all of the actions in this world but about all actions in all ideal worlds. It would surely be a mistake to go from what ought to be the case in this world to what ought to be the case in ideal worlds. But what ought to be the case in an ideal world ought also to be the case in this world, no matter how miserable this world may be.

When the quantifier is an existential one, as we have seen, the expression evaluates a particular act, but if we replace the existential quantifier with a universal quantifier, the result is an expression which can function as a principle of evaluation; a standard against which instances may be compared. Only a universally quantified expression may function as a moral rule. Since 'Oα[(x)A[x]]' tells us what α should do in every case, it is a moral rule. But 'O(∀x)α[A[x]]' does not directly tell us what α ought to do. Instead, it tells us what property all of α's deliberate actions should have without suggesting that α is the agent responsible for seeing to it that every action is an instance of 'A'. It is more like a principle or standard of evaluation than a moral rule. Unlike (22), (23) does not indicate that α is the agent responsible for seeing to it that every action is not an 'A'. So (23) and (24) are evaluative or normative principles. They differ in that (23) is an ideal principle and (24) is a practical principle. (24) is known to be true on the basis of practical experience in this world, and for all we know it will continue to summarize moral practice in this world. Furthermore, (23) and (24) are standards for perfor-
mances, whereas

\[(25) \quad 0(x)Ax\]

and

\[(26) \quad (x)0Ax\]

are an ideal principle and a practical principle, respectively, for
evaluating actions.

The third deontic category is permission. It is not, in my
opinion, a prescriptive category. The expression

\[(27) \quad P\alpha((\exists x)Ax)\]

is well-formed, but it is not a prescription since there is no clear
way that it can be consistently viewed as advocating a course of action.
Instead, permissions are claims to the effect that a contemplated action
is compatible with all of one's obligations. There is no sense in which
an ordinary permission involves imperative force. The proper way to
interpret (27) is the straight-forward way—an imperative is permis-
sible.

Looking at Table 8, we note, once again, that all roads lead
to

\[(28) \quad P(\exists x)Ax\]

which, along with

\[(29) \quad P(x)Ax\]

was defended as the proper formalization of permission. If the permis-
sion is relative to an agent who may perform the action (and, perhaps,
no one else may), then move up the chart one position to

\[(30) \quad P(\exists x)\alpha[Ax].\]
TABLE 8

AUGMENTED SQUARE OF IMPLICATION
The farther one moves against the arrows, the more stringent the claim becomes. The reader should already be aware of the difference between (30) and

\[(31) \quad Pa[\exists x A x],\]

for example.

The formal machinery of IDL is completed by making provision for the use of quantifiers over agents. All this requires is a single formation rule:

\[(FR7) \quad \text{If } A \text{ is a wff, then } (a)A \text{ is a wff.}\]

The necessary rules of inference are the standard quantification rules.

The addition of quantification over agents is not particularly exciting. The distinctions introduced so far apply equally well to expressions which involve agent quantifiers. For example,

\[(32) \quad 0(a)1a[(x)A x]\]

is a prescriptive,

\[(33) \quad (a)(x)0a[A x]\]

is a practical principle, and the difference between

\[(34) \quad \exists a 0a[p]\]

and

\[(35) \quad 0(3a)a[p]\]

should be fairly obvious. The former says that someone does exist who has the duty to see to it that 'p' is true, and the latter says someone should (but, perhaps, does not) exist who ought to see to it that 'p'.

Less clear are formulas in which act and agent quantifiers interchange. For example,
(36) \[ 0(a)(x)a[Ax] \]

is an ideal principle; a rule which must be satisfied by every good agent and also by every good act. But what of the hybrids such as

(37) \[ (a)0(x)a[Ax] \]

and

(38) \[ (x)0(a)a[Ax] \]?

The former tells us what constitutes moral behavior of moral agents in this world. The latter is quite a different matter. Principles of this form can only be regarded as extreme forms of "I am my brother's keeper." It says that given any act (no matter whose action it may be), everyone ought to see to it that it is an instance of 'A'; everything that one does is the moral concern of everyone else.

There are 32 ways to combine agent and act quantifiers (both universal and existential) with the '0' operator alone. I believe that all of them can be understood without difficulty by applying the various principles of interpretation developed in this chapter and in Chapter VI. One last point. I do not believe that the limitation on importing individual moral agents into ideal worlds must also be observed in importing individual moral agents into ideal worlds. My libertarian leanings, humanitarian hunches, and religious atavisms balk at the idea of refusing admission to anyone into the promised land. But an ideal world is not the promised land. It is a world in which all obligations are satisfied and some people are so tainted by their personal histories and conflicting personal obligations that arise from them, that even in an ideal world it would be impossible to satisfy them. Even so, the
restriction on importing act individuals is sufficient. This restriction means that even if all moral agents exist in all ideal worlds, their actions might not. The traditional bad men, Hitler, Capone, et al., may exist in any ideal world, but they might not be allowed to carry all of their actions with them. This has the pleasant consequence of reducing the number of modalities that must be investigated. Meanwhile, Table 9 summarizes the moral distinctions (the most interesting ones) that I believe can be formalized in IDL.

9.3 Dealing With Moral Exceptions

The complexity and power of IDL should be obvious, and investigating it thoroughly will take an enormous amount of time. Here, I will use it to discuss two aspects of the general problem of exceptions to moral rules. The first aspect concerns the remarkable inability of philosophers to come to grips with exceptions to moral rules, and concludes with suggestions for the use of IDL in formalizing contrary-to-duty obligations. The second aspect is the more serious problem of moral dilemmas.

In an extremely important and sophisticated paper, John Rawls distinguishes two types of rules: summary rules and constitutive rule rules\(^3\). A summary rule is simply a summary of individual decisions made in the past—a type of empirical generalization—whereas a constitutive

---

\(^3\)Rawls [1955]. Rawls speaks of "rules of practice" instead of "constitutive rules". I have borrowed the latter term from Garner and Rosen [1967], p. 87.
OPTATIVES

Evaluations of Actions

\[ \text{Ante eventum} - 0(\exists x)Ax \]
\[ \text{Post eventum} - (\exists x)0Ax \]

Evaluations of Deliberate Actions

\[ \text{Ante eventum} - 0(\exists x)\alpha[Ax] \]
\[ \text{Post eventum} - (\exists x)0\alpha[Ax] \]

OBLIGATIVES

Positive Duties - 0\alpha[(\exists x)Ax]

For all - 0(\alpha)\alpha[(\exists x)Ax]

For some - 0(\exists \alpha)\alpha[(\exists x)Ax]

Negative Duties - 0\alpha[(x)-Ax]

For all - 0(\alpha)\alpha[(x)-Ax]

For some - 0(\exists \alpha)\alpha[(x)-Ax]

PRESCRIPTIVES

Positive Duties - 0l\alpha[(\exists x)Ax]

For all - 0(\alpha)l\alpha[(\exists x)Ax]

For some - 0(\exists \alpha)l\alpha[(\exists x)Ax]

Negative Duties - 0l\alpha[(x)-Ax]

For all - 0(\alpha)l\alpha[(x)-Ax]

For some - 0(\exists \alpha)l\alpha[(x)-Ax]

MISCELLANY

Ideal Principles (Constitutive Rules) - 0(x)Ax

Constitutive Moral Rules - 0(x)\alpha[Ax]

Practical Principles (Summary Rules) - (x)0Ax

Practical Moral Rules - (x)0Ax

Prima Facie Obligations - \alpha[p \rightarrow q]

Actual Obligations - \alpha[p]
rule defines a practice. The rules of baseball, to use Rawls' example, are constitutive rules:

In a game of baseball if a batter were to ask "Can I have four strikes?" it would be assumed that he was asking what the rule was; and if, when told what the rule was, he were to say that he meant that on this occasion he thought it would be best on the whole for him to have four strikes rather than three, this would be most kindly taken as a joke.4

The type of rule whose applicability may be questioned in particular cases is a summary rule. Rawls' purpose in drawing the distinction is to avoid traditional objections to utilitarianism. These objections are individual cases in which one clearly ought not do what is best on the whole, e.g., torturing one person to eradicate disease. Such cases do not count as objections to a theory which justifies practices instead of individual actions on utilitarian grounds. To use the current jargon, the traditional objections are cogent when applied to act utilitarianism but not when applied to a rule utilitarian theory that views moral rules as constitutive rules. Constitutive rules, by their very nature, are exceptionless:

In the case of actions specified by practices it is logically impossible to perform them outside the stage-setting provided by those practices, for unless there is the practice, and unless the requisite properties are fulfilled, whatever one does, whatever movements one makes, will fail to count as a form of action which the practice specifies. What one does will be described in some other way.5

Unfortunately, Rawls spoils this crucial point in a passage which occurs


5 Rawls [1955], p. 25.
prior to the drawing of this distinction. He finishes a discussion of
promising by claiming that if a promise is broken, the defense that it
was broken because doing so was best on the whole is an unacceptable
defense because the rule "Promises ought to be kept" defines a practice
(is a constitutive rule). Even though it defines a practice, we must
allow for exceptions because

...as with any set of rules there is understood
a background of circumstances under which it is expected
to be applied and which need not--indeed which cannot--
be fully stated.

6

Rawls believes that he can say this and achieve the purpose of his
paper. Garner and Rosen, however, seize upon this unfortunate passage
as a basis for rejecting a rule utilitarian's conception of moral rules
as constitutive rules7. If all exceptions are not stated as part of
the rule, then it is always possible for there to be an exception which
is not covered by the rule which should be covered by the rule. Hence,
they conclude, no moral rule can be universally true. I have character-
ized this passage of Rawls' paper as "unfortunate" because he need not
have said it. He is right in saying that some moral rules may be viewed
as constitutive rules in the face of putative exceptions but he need not
have introduced a set of "understood circumstances" to support that
view. I agree with Garner and Rosen that this move undermines the view.

Before showing how Rawls may have exceptionless constitutive
rules, I want to follow Rosen and Garner for a moment. After discussing
rule utilitarianism, they move on to a consideration of act deontology

6 Rawls [1955], p. 17.
7 Garner and Rosen [1967], p. 77.
which they ultimately defend as the best normative theory of obligation. If I may greatly simplify their presentation, they argue that the principle

(39) If any A has properties F, G, and H, then A is right (or wrong or obligatory),

must ultimately be justified by statements like

(40) A is right (or wrong or obligatory).\(^8\)

In other words, (39) is viewed as a summary rule since, as Rawls believed he was forced to admit, one can never state all of the possible exceptions to the rule as part of the rule. Since Garner and Rosen allow us the option of replacing 'right' with 'obligatory', let us do so, and let us take as a specific example of (39) the case discussed by Rawls, keeping promises. The general principle that one ought to keep one's promises may be translated any number of ways. Two of them are\(^9\)

(41) 0\(\alpha(x) (Mx \rightarrow (\exists y) Kyx)\]

and

(42) 0\(\alpha(x) (Mx \rightarrow (\exists y) Kyx)\),

and these formulas capture a familiar difference: (41) represents a *prima facie* duty, and (42) represents an actual duty if a promise is made. Let us also suppose that there is an act of making a promise yet it is false that there ought to be an instance of keeping it (let

---

\(^8\) Garner and Rosen [1967], p. 85.

\(^9\) Reading 'Mx' as 'x is an instance of making a promise' and 'Kxy' as 'x is an instance of keeping the promise made by doing y.'
us suppose that in this case keeping the promise would result in the
deaths of thousands of innocent persons):

\[(43) \quad (\exists x)(Mx \& -0a[(\exists y)Kyx]).\]

Now, (43) and (42) are contradictories and we would describe (43) as
an exception to (42). But it is not an exception to (41). These two
are compatible. Not only is (41) a prima facie obligation but it is
a constitutive moral rule.\(^{10}\)

In ideal worlds, whatever is obligatory is true and whatever
becomes true cannot be forbidden, and that is a logical 'cannot',
otherwise the ideal world collapses into logical inconsistency.
Hence, ideal worlds define what ought to be because it is logically
impossible that anything else is part of an ideal world. If a new
action is added to an ideal world, it is logically necessary that it
be in accordance with (41). Hence (41) must be viewed as a moral
constitutive rule.

Does that mean that (42) is the form of a summary rule?
I don't think so. I believe that

\[(44) \quad (x)0a[Mx \rightarrow (\exists y)Kyx]\]

is a much better candidate to formalize the summary rule which corres-
ponds to (41). Proponents of summary rules agree that summary rules
are unscarred by putative counter-examples, cases describable by (43).
(44), unlike (42), is compatible with (43); (42) is not. (42) is an
actual obligation (provided a promise is made) and a proponent of a

---

\(^{10}\)Prima facie obligations are expressed by an '→' within the
scope of an agent operator; constitutive rules are expressed by a
universal quantifier falling within the scope of '0'.

summary rule analysis of moral rules would claim that moral rules
are not statements of actual obligations, moral rules only tell us
either what is generally morally correct or what, "as luck would
have it", has always been found to be morally correct. This accords
quite nicely with the required interpretation of (44). As we have
seen, the most that we can read into the universal quantifier is
that it is a summary of all past cases. It offers no comfort that
the future will resemble the past. So, too, in ethical matters, a
universally quantified expression which tells us something about all
actions is true only in so far as every action performed until now
does not falsify the claim made. As a predictor of what moral qual-
ties actions will have tomorrow or the next day, it is no more reli-
able than any other empirical generalization. But if that same quant-
ifier appears within the scope of a deontic operator, the story being
told is much different. Then the claim being made is that every act
in all ideal worlds has a certain character. Furthermore, ideal worlds
are not subject to the same moral corruption that we find in the actual
world. By "moral corruption" I mean the possibility of a moral gen-
eralization being falsified as time passes. Such "non-corruptable"
principles have been identified with constitutive rules.

If neither (41) nor (44) is susceptible to actual counter-
examples of the sort represented by (43), how do (41) and (44) differ?
They differ in that (44) is perfectly compatible with

(45) \[ P(3x)\alpha[Mx \& -(3y)Kyx]. \]

It is quite possible that a summary rule to the effect that all promis-
es ought to be kept (at least so far) is true, and yet it is also
the case that it is permitted (tomorrow, say) that there is a promise
which is not kept. But the corresponding constitutive rule, (41) is
not compatible with (45). If it ought to be that all promises are
kept, then it is not permitted that there be an unkept promise, i.e.,
there is no ideal world in which a promise is not kept. They also
differ in that a constitutive rule implies its corresponding summary
rule. This seems in order. Surely if a summary rule is false, the
corresponding constitutive rule must also be false. But, one might
say, cannot a constitutive rule like "All promises ought to be kept"
be true while its corresponding summary rule is false because there
are cases of promises which shouldn't be kept? This question indicates
a continued lack of appreciation of the difference between an actual
obligation, e.g., (42), and a summary rule, e.g., which has just been
described. The actual obligation to keep all promises is falsified
by this kind of case (a promise which ought not be kept), but a sum-
mary rule is only falsified if there is an action which is permitted
to be both the making of a promise and a kept promise.

It should be clear that given the nature of constitutive rules,
Rawls should not have mentioned the need for dealing with exceptions,
and it should be clear, given the logical characteristics formalizable
and exhibited by IDL, that Rawls need not have mentioned exceptions
in the same breath as constitutive rules at all. What IDL shows, of
course, is that what counts as a counter-example to a constitutive
rule is not what counts as a counter-example to a summary rule, is
not what counts as a counter-example to an actual obligation.

To Garner and Rosen, we should point out that a summary rule is a poor logical device. A glance at the Augmented Square of Implication (Table 8) should indicate, as one would expect, that no constitutive rule may be inferred from it. In response it might be said that summary rules are only rules of thumb and if they are subject to any logical laws at all it must be rule-of-thumb logic, and, at any rate, the important point for ethical theory is that rules are only justified on the basis of inductive generalizations from particular cases. Perhaps so, I do not care to choose sides in a debate that has been raging since Kant and Mill. My purpose is to clarify the logical landscape of the debate. That debate is not ended as easily as Garner and Rosen suppose. If what I have said so far has merit, the logical relations to be found among moral rules and their exceptions are more intricate than either Rawls or Garner and Rosen have indicated. Constitutive moral rules have not been done in.

There is no contradiction in supposing that "Promises ought to be kept" is a constitutive rule and also in supposing that here is a promise that ought not be kept. In fact, IDL indicates that from (41) and

\[ (46) \quad 0-(3y)a[Ky] \]

It follows that 'a' ought not be an instance of a promise, i.e.,

\[ (47) \quad 0-a[Ma]. \]

This seems to accord quite nicely with conventional moral wisdom. But the temptation to say "Big deal. The promise has already been made."
Now what do I do?" cannot be resisted.

By way of preliminary remarks in answering that question, I would like to point out that the answer is not a matter of logic—it is a genuine ethical concern. It is not just a matter of what constitutes consistency among propositions containing ethical terms, it is a matter of what conditions must be true in order for propositions involving ethical terms to be true. In continuing the discussion it will be advantageous to do two things: drop quantifiers and use CL notation.

In the following, 'p' will be read as 'α promises' and 'q' will stand for 'α keeps his promise.' The prima facie obligation of α to keep his promises is represented by

\[(48) \quad 0α[p \rightarrow q].\]

But we are supposing that in this case it is forbidden that α keep his promise:

\[(49) \quad 0α[\neg p].\]

What is needed at this point is a moral principle that tells what to do if we do not keep a promise. Let us suppose that one ought to say 'I'm sorry' when a promise isn't kept. As a prima facie obligation this principle would look like

\[(50) \quad 0α[\neg q \rightarrow r].\]

From (48), (49), and (50), two things follow. First, that α ought not have promised (which we already knew), and, second, that α ought to say 'I'm sorry'. To know what to do when one has done something which is forbidden requires more than just knowing deontic logic, it also
requires knowing substantive moral principles. Given those principles, deontic logic (IDL) can be used to show us what does and what does not follow from them.

The situation just discussed is not a contrary-to-duty case of the sort discussed in the closing moments of Chapter VI. The reason being that α's obligation to keep his promise was a *prima facie* obligation, so (50) is not, strictly speaking, a contrary-to-duty obligation. And it is now time to confess that the solution offered at the close of Chapter VI to the contrary-to-duty paradox is not entirely successful. Let us suppose for a moment that α has an actual obligation to never lie:

\[(51) \quad \neg \alpha(x) \land \phi(x)\]

(I have reverted to quantifiers and I am not indicating agency so that I can show that quantifiers alone do not solve the paradox.) From (51) we may derive

\[(52) \quad \forall \alpha(x)(\forall y \phi(x) \land \exists y \psi(y)) \]

which says that every act of lying should be followed by an act of stealing. By indicating agency (*via* CL), this undesirable result may be avoided. The obligation to never lie may be formalized as

\[(53) \quad \neg \alpha(x) \land \phi(x) \]

and nothing which says that α has an obligation to steal, if he lies, may be derived from (53). As far as I can tell, IDL handles contrary-to-duty obligations in an entirely satisfactory way. Moral dilemmas, however, are another matter.
Moral dilemmas are a bit of a dilemma themselves. First, it is not entirely clear that there are such things. Second, this is partly because it isn't clear what their nature is. IDL can be of some use here, but not much. In the first place, 

(54) \[Op \& Op\]

is not a moral dilemma. It is a deontic contradiction, hence, the state of affairs represented by (54) is either logically or conceptually impossible, or else, at least, analytically false just as 'This is a red and blue (all over) book' is analytically false. If a standard deontic logic is going to be capable of dealing with moral dilemmas, then a moral dilemma must be expressed in some way other than (54). Walzer provides a happy suggestion. A person is faced with a moral dilemma, he says, if a person is faced with "a situation where he must choose between two courses of action both of which it would be wrong for him to undertake." Since (54) is ruled out we must formalize the claim that both courses of action are forbidden as

(55) \[0a[-p] \& 0a[-q],\]

but must choose between 'p' and 'q'. Hence,

(56) \[La[p v q]\]

must be conjoined to (55). This conjunction represents a moral dilemma. It is fortunate that the ought-implies-can principle was ruled out in Chapter VI as a logical consequence, otherwise the conjunction of (55) and (56) would be inconsistent and any principle telling us

\[11^{11}\text{Walzer [1973], p. 160.}\]

\[12^{12}\text{Ibid.}\]
what to do in the case of a moral dilemma would be a tautology since its antecedent would be a contradiction. Hence, such a principle would tell us nothing. Next, note that

\[(57) \quad 0((0a[-p] \& 0a[-q] \& La[p \lor q]) \rightarrow r)\]

is a deontic tautology because in an ideal world the antecedent of the conditional is inconsistent. Apparently, we must choose (57) or

\[(58) \quad (0a[-p] \& 0a[-q] \& La[p \lor q]) \lor 0r)\]

as the appropriate forms of moral principles that tell us what to do in the event of a moral dilemma. (57), as we have seen, is unacceptable which means that principles which are relevant to resolving moral dilemmas cannot be constitutive moral principles nor can they express *prima facie* obligations. How can that be explained? The essence of a moral dilemma is that however the dilemma is resolved, something wrong must be done. It makes sense that no constitutive rule could advocate that something wrong be done, which it would have to do. And no one could have a *prima facie* duty to do something which is assumed to be wrong, e.g., murder, whereas one may have an actual duty to do something which is *prima facie* wrong. In the case of a moral dilemma we are faced with a situation where we must do something wrong. That leaves (58). But it fares no better. The consequent of (58) must include either '0a[p]' or '0a[q]' which, in effect, tells us to do something wrong; that 'a[p]', for example, is true in all ideal worlds. But that is logically impossible since the antecedent tells us that 'a[-p]' is true in all ideal worlds. According to the semantic basis for IDL, if the antecedent of (58) is true, its consequent
is false, and if (58) is true, then its antecedent is false. In other words, (58) is equivalent to the denial of its antecedent. IDL, it would seem, is incapable of formalizing the arguments one would use to deal with moral dilemmas. Another operator is required, perhaps a utility operator, 'Up', which might be read as 'p is not ideal but it is best on the whole.' This strikes me as being the next most important task of deontic logic—to work out the logic of 'best-on-the-whole' which might throw some more light on the relationships between constitutive moral rules and principles of utility which, hopefully, would throw some light on the debate between Kant and Hill alluded to earlier.

This ends my examination of monadic deontic logic. We have looked at the major historical developments in an attempt to cull the valuable insights from the limitations and errors of the past, and turn them into a deontic logic which represents the best that monadic deontic logic based on an ideal world semantics can become. We terminated the investigation by examining an important limitation of IDL. But that limitation does not mean that IDL or deontic logic should be given up. That mistake has been made too many times in the history of deontic logic.
BIBLIOGRAPHY

Anderson, Alan Ross

Anderson, Alan Ross and Moore, Omar Khayyam

Ayer, A. J.

Barnes, W. H. F.

Bohnert, H. G.

Brandt, Richard B.

Castaneda, Hector-Nerl
Chisholm, Roderick M.

Copi, Irving M.

Espersen, Jon

Føllesdal, Dagfinn and Hilpinen, Risto

Garner, Richard T. and Rosen, Bernard

Gombay, Andre

Grelling, Kurt

Hansson, Bengt

Hedenius, Ingemar
[1941] *Om rätt och moral* [On Law and Morals], Stockholm, 1941.

Hilpinen, Risto

Hintikka, Jaakko

Hughes, G. E. and Cresswell, M. J.

Jeffry, Richard C.
Jørgensen, Jørgen  

Kalinowski, Jerzy  

Keene, C. B.  

Kielkopf, Charles F.  

Lemmon, E. J. and Nowell-Smith, P. H.  

Lewis, Clarence I.  

Lewis, Clarence I. and Langford, Cooper H.  

Mally, Ernst  

Meredith, David  

Moritz, H.  

Prior, A. N.  


Purtill, Richard L.  
references are to a copy of the paper which the author kindly sent to me.

Rawls, John

Rescher, Nicholas

Ross, Alf

Searle, John

Smiley, T. J.

Sosa, E.

Stevenson, C. L.

Taylor, Paul

Walzer, Michael

Williams, B. A. O.
von Wright, Georg Henrik


