Karl Pettersson

The Logical Structure of the Moral Concepts

An Essay in Propositional Deontic Logic

UPPSALA UNIVERSITY
Dissertation presented at Uppsala University to be publicly examined in Geijersalen, Engelska parken, Thunbergsvägen 3P, Uppsala, Saturday, January 29, 2011 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in Swedish.

Abstract

In this thesis, the main focus is on deontic logic as a tool for formal representation of moral reasoning in natural language. The simple standard system of deontic logic (SDL), i.e. the minimal Kripkean modal logic extended with the deontic axiom, stating that necessity (interpreted as obligation) implies possibility (interpreted as permission), has often been considered inadequate for this aim, due to different problems, e.g. the so-called deontic paradoxes. A general survey of deontic logic and the problems with SDL is made in chapter 1. In chapter 2, a system denoted Classical Deontic-Modal logic (CDM1) is defined. In this system, there is a primary obligation operator indexed to sets of possible worlds, and a secondary requirement operator, defined in terms of strictly necessary conditions for fulfilling an obligation. This secondary operator has most of the properties of the necessity operator in SDL. In chapters 3 and 4, it is argued that CDM1 is able to handle the SDL problems presented in chapter 1 in an adequate way, and the treatment of these problems in CDM1 is also compared with their treatment in some other well-known deontic systems. In chapter 5, it is argued that even though the problems related to quantification in modal contexts are relevant to deontic logic, these issues are not specific to deontic logic. In chapter 6, the relations between some controversial features of moral reasoning, such as moral dilemmas and “non-standard” deontic categories like supererogation, and deontic logic are discussed. It is shown how CDM1 can be modified in order to accommodate these features.

Keywords: deontic logic, standard deontic logic (SDL), deontic paradoxes, non-Kripkean modal logic, moral dilemmas, supererogation, Hector-Neri Castañeda, practitioners, dyadic deontic logic

Karl Pettersson, Department of Philosophy, Ethics and Social Philosophy, Box 627, Uppsala University, SE-75126 Uppsala, Sweden.

© Karl Pettersson 2010


urn:nbn:se:uu:diva-131581 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-131581)

Printed in Sweden by Universitetstryckeriet, Uppsala 2010.
4.2.1 The Alf Ross Paradox in CDM1 .................................. 45
4.2.2 The Consensus about the Alf Ross Paradox .................. 46
4.2.3 Danielsson’s Arguments Against the Ross Inference ....... 46
4.3 The Good Samaritan Paradoxes ..................................... 49
4.3.1 The Good Samaritan Paradoxes in CDM1 ................... 49
4.3.2 The Scope Solution in SDL .................................... 50
4.3.3 Castañeda’s System ........................................... 52
4.3.4 Lewis-style Systems .......................................... 53
4.4 Contrary-to-Duty Paradoxes ........................................ 54
4.4.1 Contrary-to-Duty Paradoxes in CDM1 ....................... 54
4.4.2 Castañeda’s System ........................................... 56
4.4.3 Lewis-style Systems .......................................... 57
4.4.4 Danielsson on Contrary-to-Duty Paradoxes ............... 59
4.5 Geach’s John/Tom-example ......................................... 60
5 Quantification in Deontic Logic ...................................... 63
5.1 Introduction ...................................................... 63
5.2 General Issues in Quantified Modal Logic ..................... 63
5.3 Extensionality as a Case for Special Treatment of Quantified Deontic Logic .............................................. 64
5.3.1 Castañeda’s Extensionality Argument ....................... 64
5.3.2 Substitution of Identical Individuals and the Jephta Dilemma ....................................................... 65
5.3.3 Forrester and the Multiplied Beneficence Argument .... 66
6 Supererogation and Dilemmas ........................................ 69
6.1 Introduction ...................................................... 69
6.2 Chisholm on Supererogation ...................................... 69
6.2.1 Chisholm’s Nine-Category Proposal ....................... 69
6.2.2 Chisholm’s Agent-Relative Proposal ....................... 72
6.3 McNamara’s View on Supererogation ......................... 73
6.4 Deontic Dilemmas and Classic Logic ............................ 75
6.4.1 Schotch’s and Jennings’s Non-Kripkean Deontic Logic . 75
6.4.2 Goble’s Proposal ............................................. 76
6.5 A Relevant Approach: Routley’s and Plumwood’s Deontic Logic ..................................................... 77
6.6 Danielsson on Prescriptive and Ascriptive Norms .......... 79
6.7 Incorporating a Solution into CDM ............................... 80
6.7.1 Dilemmas in the CDM Framework ......................... 80
6.7.2 A DWE-like Assessment of Supererogation ........... 83
7 Appendix .......................................................... 87
7.1 SDL Formulae ................................................... 87
7.2 CDM1 Formulae .................................................. 88
7.3 Comparison Table for Systems .................................. 90
Bibliography ................................................................ 93
Index ......................................................................... 97
1. Introduction

1.1 The Plan of this Thesis

The aim of this thesis is to discuss the possibility of representing logical rules governing normative reasoning, in particular moral reasoning, in ordinary language, by means of coherent, formal systems. I will discuss different challenges to this type of representation, and see whether they can be met.

In this chapter, I will briefly explain the basic idea of deontic logic, and make a survey of some of the problems that have been considered most important in the debate surrounding it. I will discuss possible solutions to most of these problems in later chapters. However, in this chapter, I will critically discuss some general problems relevant to the structure and methodology of this thesis, as well as the possibility of applications of deontic logic in other areas of philosophy. I will assume that the reader is familiar with standard propositional logic as well as with some basic concepts of metalogics and modal logic, such as the concept of an axiom or a derivation rule.

In chapter 2, I will present a system of deontic logic, denoted CDM1 (with the abbreviation CDM standing for Classical Deontic-Modal logic). In chapter 3, I will give a better idea of the distinguishing features of this system by comparing it to some other systems that have been proposed by different authors. In chapter 4, I will discuss how CDM1 handles the problems I present in the introduction compared to the other systems discussed in chapter 3. In chapter 5, I will discuss some issues about quantification in deontic logic, and defend the claim that the controversial issues in this area are not specific to deontic logic, which is the reason why I do not want to defend any specific deontic logic with quantification over ordinary individuals (CDM1 does contain quantification over propositions). In chapter 6, I will briefly discuss how some “non-standard” deontic concepts, like dilemmas and supererogation might be handled in deontic logic.

1.2 General Conventions

Generally, I use standard predicate calculus symbolism, but the following conventions may be worth noting: Hector-Neri Castañeda uses small letters to refer to propositions, and capital letters to refer to practitioners, a practice I will follow in this thesis.
1.3 Methodological Issues

In this thesis, I will claim the aim of deontic logic to roughly consist in providing a theoretical systematization of moral reasoning in ordinary language. I will not say much about possible differences between moral reasoning and normative reasoning in other contexts, e.g. in the context of law\(^1\).

If we cannot find a translation of natural moral language into the formal language of a deontic system without either justifying an inference pattern that ordinary reasoning sees as invalid, or failing to justify an inference pattern that ordinary reasoning sees as valid, and if this cannot plausibly be explained by appealing to pragmatic rules of communication or some kind of reliance on implicit contingent premises, this would speak against the system.

Lennart Åqvist speaks of a translation \( t \) of natural language into the formal language of a system of deontic logic \( \mathcal{L} \) as being fully adequate with respect to deontic reasoning in natural English (NDL) just in case any English sentence is valid in NDL if and only if its \( t \)-translation is valid in \( \mathcal{L} \) [3, p.40]. Ideally, my goal with this dissertation would consist in formulating a system for which we can find a fully adequate translation from NDL into the formal language of the system.

It is, of course, not possible to the test a deontic system against every possible inference. In this dissertation, I will focus on discussing some types of situations, where it is acknowledged that the simple so-called standard system of deontic logic SDL, at least prima facie, has problems representing inference patterns in ordinary language. SDL and its problems will be introduced later in this chapter. The system CDM1 will also, in at least one aspect, be too weak in its expressive power to find a fully adequate translation in Åqvist’s sense from NDL into CDM1, because it, for the reasons developed in chapter 5, contains no quantification over ordinary individuals. One could argue that a fully adequate system also should contain e.g. logical principles for other intensional contexts, such as beliefs, because they may be relevant to normative reasoning. My hope is, however, that CDM1 may be extended into a system into which we could find a fully adequate translation from NDL, if we find satisfactory solutions for the various problems in other areas of logic not directly related to the logic of normative concepts.

We should not expect our deontic system to exactly mirror the inference patterns of ordinary language. It is very plausible to claim that ordinary speakers have no definite intuitions about the validity of certain inference patterns in many particular cases, especially when considering very complicated or abstract statements. Of course, people might also make mistakes in their reasoning. We can expect situations where e.g. inferences that speakers prima facie are prone to make come out as invalid in our system. We should understand

\(^1\)I do not even want to exclude the possibility that there exist different uses of normative expressions within the realm of moral language: my goals for this dissertation are satisfied if there is one central use for which I have provided a systematization.
the aim of giving a model of moral reasoning in natural language so that the system must still fit the inference patterns in the sense that speakers would be willing to correct their inferences in accordance with the system if they were able to understand how they were made.

If we have different deontic systems that, on the whole, fit the patterns of ordinary reasoning equally well, other considerations, such as theoretical simplicity, and keeping with philosophical tradition, come into play. It speaks in favor of a deontic system if it does not deviate too much from the tradition of philosophical logic, e.g. if it makes use of semantic structures and rules similar to Kripkean modal logic. This makes the system easier to understand and discuss in the philosophical community. I will define CDM1 as a system based on a sort of modal logic, with the deontic expressions interpreted as propositional operators. I will not go very much into modern systems with non-propositional operators, where deontic logic often is connected with the logic of action and intention, even though I will compare CDM1 with some such systems, e.g. Castañeda’s, as to how the SDL problems are handled in the different systems.

A problematic situation would be if people had definite but conflicting intuitions about the validity of certain inference patterns in normative reasoning, and if this could not be explained by appealing to pragmatic rules or different underlying factual or substantive normative assumptions. In such cases, it may be that different people really are using the normative words in different senses, and we have to make a choice about which set of intuitions we should consider relevant to the justification of the system. The possibility of genuine moral dilemmas or genuine supererogation may be cases in point. A full investigation into these matters would require a meta-ethical discussion beyond the scope of this thesis. My main proposal of a deontic system in chapter 2 will be defined so that it excludes both dilemmas and supererogation. However, in chapter 6, I will discuss how the system might be modified in order to accommodate these features.

1.4 Basic Deontic Logic

Deontic logic is supposed to be about what statements of norms logically imply, and what they are implied by. If we accept a normative sentence like “it is obligatory to \( \phi \)”, “you ought to do \( \phi \)”, or “\( \phi \) is forbidden”, which other sentences do we have to accept if we are not to be charged with inconsistency? Just as ordinary propositional or predicate calculus tries to capture the logical relations between statements, or propositions in general, by giving general axioms or derivation rules, so does deontic logic have its own axioms or rules that specify logical relations between normative statements, as well as between normative statements and statements in general. If we can find some instance of valid reasoning where we need to perform some oper-
ations on the things to which normative expressions, like “ought”, apply, we have at least a reason for going beyond a simple first-order logic treatment, where Ought(x) is treated as a predicate that applies to non-complex individual terms\(^2\). It seems that such instances can be found. It seems consistent to claim: “You ought to buy tea or coffee, but it is not the case that you ought to buy tea, and it is not the case that you ought to buy coffee”. We could try the following representation in first-order logic, where \(O(x)\) represents the “ought”-predicate, and \(t\) and \(c\) represent some individuals consisting in something like the actions of buying tea and coffee.

\[
\text{(Buy)} \quad (O(t) \lor O(c)) \land (\neg O(t) \land \neg O(c))
\]

(Buy) is obviously self-contradictory. Apparently, we have one obligation to buy tea or coffee. We can treat this entity that the obligation seems to operate on, buying-tea-or-coffee, as one individual, as in:

\[
\text{(Buy*)} \quad O(t_\text{or}_c) \land (\neg O(t) \land \neg O(c))
\]

(Buy*) is not self-contradictory in first-order logic. But many intuitively valid inferences will be invalid. “You ought to buy tea or coffee” seems to be logically equivalent to e.g. “You ought not to buy neither tea nor coffee” and “You ought to buy tea if you do not buy coffee, and to buy coffee if you do not buy tea”. How can we express these equivalences, if we do not allow “ought” to operate on logically complex entities?

Every logical system that has been proposed as a system of deontic logic contains an operator \(\Box(\phi)\) which says that something, \(\phi\), is obligatory. What kind of entity \(\phi\) stands for, and, consequently, exactly how \(\Box(\phi)\) is to be interpreted, varies from system to system. In most systems, \(\phi\) stands for some proposition, and \(\Box(\phi)\) should be read as “it is obligatory that \(\phi\) is the case” or “it ought to be that \(\phi\) is the case”.

However, in some systems \(\phi\) stands for a name of an act, and \(\Box(\phi)\) should be read as “it is obligatory to do \(\phi\)” or “\(\phi\) ought to be done”. One of the contested issues in deontic logic is whether there is any fundamental difference between statements about what ought to be the case and statements about what agents ought to do. Proponents of the systems where \(\Box\) operates on propositions usually say that a statement like “You ought to prepare for your seminar” can be translated into something like “It ought to be the case that you prepare for your seminar”, but those who have other alternatives, e.g. Hector-Neri

\^[2\]This does not mean that a first-order language is in itself inadequate for deontic logic. Modal claims can be expressed in first-order logic, with quantification over worlds, and some philosophers would like to treat modality as a predicate of sentences (with additional axioms, but within the framework of first-order logic), and not as a sentential operator. If this is the right approach for modality in general, it should perhaps also be applied to deontic inferences. (As we soon will see, deontic logic has often been considered a kind of modal logic.) Nevertheless, some work in specifying the correct deontic axioms would remain.
Castañeda in *Thinking and Doing* [9] and other works, may say that none of the two kinds of oughts can be analyzed in terms of the other. In many systems the \( \Diamond \)-operator is also relativized to times, agents or obligation systems. Much of this is done in order to find a way out of the difficulties with the simple so-called standard system, which I will mention later in this chapter.

“Ought” is a modal verb like, for example, “must”, “can”, “knows”. In the beginning of the article that may be said to have started modern debate on deontic logic, Georg Henrik von Wright pointed out the analogy between deontic logic and the logic of other modalities, such as so-called alethic modal logic, which is about what is, in a logical or metaphysical sense, necessarily or possibly true, or epistemic modality, which is about what is known or unknown [61, p.1]. Obligation has been interpreted as “deontic necessity”, and permission has been interpreted as “deontic possibility”, and these concepts seem to be analogous, but not equivalent to alethic or epistemic necessity and possibility. The debate on deontic logic has since then focused on this analogy.

A system that has been called “the standard system of deontic logic” (SDL) (for the first specification, see [31]) seems to have been accepted as a starting-point by most writers on deontic logic; it is a proper subsystem of many systems of deontic logic. Even when it is not, as when the deontic operators are interpreted as operating on something like act-predicates rather than, as in SDL and probably most other systems, on propositions, most authors seem to have wanted to include some adapted form of its theorems. SDL is the minimal Kripkean modal logic (K) with the obligation operator \( \Box \) substituted for the standard modal necessity operator \( \Box \), and a permissibility operator \( P \) substituted for the possibility operator \( \Diamond \), i.e. \( P(a) \) is defined as \( \neg \Box (\neg a) \), and extended with an axiom which says that if \( a \) is obligatory, then \( a \) is permitted. A difference between Hansson’s original formulation and standard modal logic is that Hansson does not allow for iterated deontic operators, like \( \Box (\Box (a)) \). He does so in order to make the system more similar to action-operator systems, like von Wright’s, where such sentences are not considered well-formed. Sometimes, an operator, \( F(a) \), which is to be read as saying that \( a \) is forbidden, is defined as \( \Box (\neg a) \): what is forbidden is what ought not to be the case. Formally, SDL can be specified as a modal system that contains the “Kripke axiom”

\[
(K) \vdash \Box (a \supset b) \supset (\Box (a) \supset \Box (b))
\]

and the “deontic axiom”

\[
(D) \vdash \Box (a) \supset P(a)
\]

and is closed under the derivation rule (RN), “rule of necessity”, stating that any theorem in the underlying logic has obligatory status³

³In Bengt Hansson’s system, which does not allow iterated operators, this means that \( a \) is a theorem in the underlying basic logic (e.g. propositional logic). In standard modal logic, however, this means that \( a \) is any theorem in the modal system.
(K) and (RN) together make SDL into a so-called normal modal system, which can be given a standard Kripkean semantics, where $\Box(a)$ means “$a$ is true in every accessible (which is here often interpreted as ‘permissible’) world”. The axiom schema (D) may also seem intuitively very plausible: if something is obligatory, it is also permissible. (In terms of Kripkean semantics, (D) adds that, for every world, there is at least one permissible world.)

As can easily be seen, (K) and (RN) together entail the so-called “rule of regularity” (RR), which in these contexts often is referred to as “the consequence principle”. As we will see later, the implications of this principle have been the focus of much of the debate about the paradoxes in SDL.

$\vdash a \supset b$ $\vdash \Box(a) \supset \Box(b)$

An alternative characteristic of SDL, which is more common among writers on deontic logic, is using the axiom (D), but replacing (K) and (RN) with the rule (RE), “rule of extensionality”

$\vdash a \leftrightarrow b$ $\vdash \Box(a) \leftrightarrow \Box(b)$

and the axioms of conjunctive deontic elimination and introduction, as well as the law of the excluded deontic middle

$\vdash \Box(a \land b) \supset \Box(a)$

$\vdash \Box(a) \land \Box(b) \supset \Box(a \land b)$

$\vdash \Box(a \lor \neg a)$

(EMd) is a theorem in any normal modal system. $\Box(a \lor \neg a)$ is also entailed by $\Box(a)$ if we assume the other axioms.

It seems that the analogy between deontic logic and alethic or epistemic modal logic cannot be brought much further than it is with SDL without getting into serious trouble. The concept of logical necessity is often thought to be captured by C.I. Lewis’s fifth system (S5), the deontic counterpart of which we would obtain if we replaced (D) in SDL with the stronger “truth axiom”
and added the “Brouwer axiom” (named after the Dutch mathematician L.E.J. Brouwer)

(B) $\vdash P \Box(a) \supset a$

and the axiom also used in Lewis’s fourth system

(S4) $\vdash \Box(a) \supset \Box \Box(a)$

But (T), which also seems necessary in a system of epistemic logic where $\Box(a)$ is interpreted as “$a$ is known”, and (B) are both obviously implausible in a system of deontic logic in which $\Box(a)$ means that $a$ is something that, in a normative sense, ought to be the case. Propositions with iterated deontic modalities, like any instance of (S4) (or (B)) are, as said above, not considered well-formed in systems like von Wright’s, where the obligation operator operates on act-predicates.

Do iterated deontic statements exist in natural language? It is hard to deny that we sometimes make it the case that we, or other persons, have certain obligations, and that these obligation-creating acts may well be subject to moral or other normative evaluation, such as our act of hurting a person, which makes it the case that we ought to apologize. We may thus say things like “You should act so that you have nothing to apologize for”, which can be interpreted as at least implying “It is obligatory that it is not obligatory that you apologize”. If we let $a$ stand for “You apologize”, we might want to formalize this as $\Box(\neg \Box(a))$ or $\Box(\Box(\neg a))$. In a late paper, von Wright discusses normative sentences expressing that a legal authority may, or may not, be permitted by a higher-order authority to make certain things obligatory under a norm. Such moral or legal sentences do not seem very unnatural, and it may be tempting to interpret this as a case of a deontic statement with iterated norms [65, p. 7f].

At least institutional deontic statements (though perhaps not moral statements) may be objects of knowledge, and propositions of knowledge may be objects of deontic statements. Consider a statement like “You should know that you ought to keep promises”. If we let $k$ stand for the statement that you keep promises and $\text{JB}_p(a)$ for a statement of the form that $p$ has a justified belief that $a$, it seems natural to formalize the deontic statement as $\Box(\Box(k) \land \text{JB}_p(\Box(k)))$.

1.5 Problems with SDL

Few, if any, writers on deontic logic have found SDL acceptable as a system that allows us to make all the inferences we want to make to and from normative statements, without being too strong and permitting unwanted inferences. In this section, I will go through some of the most important aspects in which SDL has been considered faulty.
1.5.1 Problems Arising from General Applications of (RR)

1.5.1.1 The Alf Ross Paradox

In the 1930s, some philosophers proposed theories of the logic of imperatives, which interpret logical relations between imperatives as being between the satisfaction conditions of the imperatives. That an imperative $J$ follows from an imperative $I$ may then roughly mean that $J$, by logical necessity, is successfully obeyed if $I$ is. Alf Ross thought that such theories cannot capture our intuitive sense of logical relations between imperatives, because they allow unintuitive inferences like:

1. Slip the letter into the letter-box.
2. $\therefore$ Either slip the letter into the letter-box or burn it [48, p.40].

This observation can easily be transferred to SDL. $a \supset a \lor b$ is a theorem in ordinary logic. So it follows from (RR) that the following principle is a theorem in SDL.

$$(\lor I_d) \circ (a) \supset \circ (a \lor b)$$

But it seems odd to derive for example “The letter ought to be burnt or posted” from “The letter ought to be posted”. The defender of (RR) has to explain this oddity somehow.

1.5.1.2 The Good Samaritan Paradox

If we have a statement expressing that it ought to be the case that a robbed person is helped, it may seem natural to formalize this as $\circ (a)$, where $a$ stands for “the robbed person is helped”. But $a$ logically implies that there exists a robbed person. With (RR), $\circ (a)$ then entails that there ought to be a robbed person, which seems unintuitive. This paradox seems to have been first formulated by A.N. Prior in [47, p. 144], but there exists several well-known versions of it, like “the Paradox of the Knower”, first formulated by Lennart Åqvist [2, p. 367]. Say that we have “You ought to know that Smith robs Jones”. In accordance with the classical analysis of knowledge, we can analyze this as $\circ (r \land JB_{you}(r))$, where $r$ stands for “Smith robs Jones”, and $JB_p(a)$ for the statement that $p$ has a justified belief that $a$. We can then, because $\circ (a \land b) \supset \circ (a)$ is a theorem according to (RR), derive $\circ (r)$, viz. that it ought to be that Smith robs Jones, which we do not want.

1.5.2 Problems Related to Conditional Obligations

1.5.2.1 Conditional Obligations in General

Arguments with conditional obligations are very common in natural language reasoning. For example, consider the following type of argument:

1. It is obligatory that $b$, if $a$.
2. $a$
3. Therefore, it is obligatory that $b$.

How is this to be represented in SDL? There are some alleged paradoxes described by Prior in [46] related to the use of (RR) in conjunction with the use of material conditional to represent implication. These may in part be viewed as transformations to the context of deontic logic of some of the usual arguments against using the material conditional as a representation of our ordinary concept of implication. It is provable in propositional logic that $a \supset (b \supset a)$, and that $\neg a \supset (a \supset b)$. Hence, it is provable in SDL that $\Box(a) \supset \Box(b \supset a)$, and $\Box(\neg a) \supset \Box(a \supset b)$. If we read $a \supset b$ as “if $a$, then $b$”, we must say that, in SDL, if something is obligatory, we are committed to it by anything, and that if something is forbidden, it commits us to anything. It might seem prima facie plausible to use a simple schema of the following type to handle arguments with conditional obligations:

$(FD) \vdash \Box(a \supset b) \land a \supset \Box(b)$

However, (FD) is not valid in SDL, and it would also have catastrophic consequences in conjunction with (RR), like in the argument below:

1. $\Box(a), P$
2. $\neg a, P$
3. $\Box(\neg a \supset b), RR, 1$
4. $\Box(b), FD, 2, 4$

1.5.2.2 The Chisholm Paradox

This paradox, which was first formulated by Roderick Chisholm in 1963, in [13], is about what one ought to do when one fails to fulfill some obligations, as in the following premises:

1. Jones ought to go to visit Smith.
2. Jones ought to notify Smith about visiting if Jones is to visit Smith.
3. If Jones is not to visit Smith, Jones ought not to notify Smith.
4. Jones is not to visit Smith.

These premises do not seem to be inconsistent or to give rise to any special confusion about what to do. But how are they to be represented in SDL? The material conditional is the only resource we have to represent conditionals.

Consider the following. Let $g$ stand for “Jones goes to visit Smith” and $n$ stand for “Jones notifies Smith about visiting”.

1. $\Box(g)$
2. $\Box(g \supset n)$
3. $\neg g \supset \Box(\neg n)$
4. $\neg g$

---

4We may note that a puzzle with a structure very similar to the Chisholm Paradox was described by Lennart Åqvist in 1960, in [1, p. 148], where it is attributed to Thorild Dahlquist.
From these premises, we can derive both $\Box(n)$ (from the first and the second, by application of (K)), and $\Box(\neg n)$ (from the third and the fourth, by modus ponens), which give rise to a moral dilemma.

Consider this alternative interpretation of the four premises:

1. $\Box(g)$
2. $g \supset \Box(n)$
3. $\neg g \supset \Box(\neg n)$
4. $\neg g$

With this interpretation, no dilemma can be derived. However, the second premise is entailed by the fourth, due to the well-known general properties of material implication, though they should intuitively be logically independent. We may note that we can just as well derive, for example, $g \supset \Box(\neg n)$. It may seem sensible to assert the second premise even though we are convinced that $\neg g$ is true, while we do not normally assert an indicative conditional when we are convinced that its antecedent is false. For that reason, this cannot be easily dismissed as just a case of the general problems with material implication as an interpretation of indicative conditionals.

1.5.3 (RE) and Geach’s John/Tom-example

The following statements seem to be analytically equivalent:

(BeatActive) John beats Tom up.

(BeatPassive) Tom is beaten up by John.

A prima facie natural SDL formalization of a statement expressing that John ought to beat Tom up would be $\Box$(BeatActive). If we hold that $\Box(a)$ should imply $\Box(b)$ if $a$ and $b$ are analytically equivalent, $\Box$(BeatActive) implies $\Box$(BeatPassive). However, as Peter Geach has pointed out, it does not seem inconsistent to claim that John ought not to beat Tom up, while Tom ought to be beaten up by John [27, p. 3]. This example may perhaps seem more plausible if we switch the roles so that John ought to beat Tom, while it is not the case that Tom ought to be beaten by John: if the two are involved in some sort of fight, we might claim that John should beat Tom, even if we do not want to claim Tom has any obligation to be beaten up.

1.5.4 Alleged Problems with the Ought-to-Be Reduction

If SDL is to function as a theory of moral ought, we must find a way to reduce statements about what agents ought to do to ought-statements operating on propositions. The simplest way might be that mentioned in section 1.4: to interpret statements of the form “agent A ought to perform action $\phi$” as $\Box(A$ performs $\phi$). Consider the following pair of statements:
It ought not to be that the child works in a factory.

The child ought not to work in the factory.

Examples of this type have been discussed by Gilbert Harman in [33, p. 131]. We might accept (FactoryBe), but decline (FactoryDo); we perhaps do not want to say that the child ought not to work in the factory, because we think there is not much the child can do about it. This may be a good objection against the universal validity of the simple type of reduction mentioned above, but it does not seem to be a serious objection against the reduction of statements about what agents ought to do to propositional ought-statements itself. We can, as e.g. John Hory has pointed out in [35, p. 53], use a more sophisticated analysis of statements like (FactoryDo), or like ⊢(the child sees to it that it does not work in a factory). I will not consider this type of objection any further.

John Hory himself considers it a problem with the reduction of ought-to-do that it becomes too tied to a utilitarian conception of morality [35, p. 53–54], but this is because he applies some quasi-utilitarian interpretation of ought-to-do. There is nothing in the semantic framework of SDL itself that forces us to use such an interpretation (for a similar point, see David Lewis’s comment in [40]. However, one might consider it a problem that the semantic framework of SDL is agent- and context-neutral, in the sense that all ought-statements are evaluated relative to the same set of ‘ideal’ worlds. This problem might also be reflected in e.g. the Good Samaritan and Chisholm paradoxes as well as Geach’s John/Tom-example. The deontic system I define in chapter 2 allows for multiple contexts, and I will discuss this in relation to the mentioned problems. I will not consider Hory’s objection as a separate problem any further.

James W. Forrester’s arguments for a fundamental distinction between the propositional ought-to-be and the ought-to-do of actions in [26, p. 55–77] do not seem to contain many arguments not related to Harman’s points or to the other SDL problems discussed in this chapters. Whether there are any advantages with a non-propositional approach in connection with the other SDL problems will be discussed in chapters 3–4.

1.5.5 The Plausibility of (D)

The axiom (D), which says that there cannot (without yielding a propositional-level contradiction of the form ⊢(a) ∧ ¬ ⊢(a)) be a proposition a such that both a and its negation are obligatory, might seem very plausible. It might be held that it captures something essential to the action-guiding nature of morality: that we cannot in this way be given contradictory obligations is something that follows from some relevant interpretation of the old slogan that ought implies can. However, (D) has been the subject of some controversy. Are there not possible situations when we face genuine dilemmas, when we have two obligations that cannot be jointly satisfied? Examples of situations
where some might hold that we really face such dilemmas are when we have to choose between sacrificing different people, or when we have made incompatible promises to different people.

In this context, we may distinguish between two kinds of impossibility to fulfill all of one’s obligations. In some cases, we have a straightforward logical contradiction between obligations, e.g. if we assume that one ought to do everything one has promised, and one has made logically incompatible promises to do both $a$ and $\neg a$, it seems to follow that $\Diamond (a) \land \Diamond (\neg a)$, which directly contradicts (D); there are no possible worlds where both $a$ and $\neg a$ are true. In other situations, there may be a causal impossibility to fulfill all one’s obligations, or to avoid all prohibited states. An example of this might be situations like the novel *Sophie’s Choice* [56], where the Auschwitz prisoner Sophie has to choose between sacrificing any of her two children. Let $c_1$ and $c_2$ stand for propositions stating that the different children are sacrificed. The statement that both of these are forbidden may be formalized as $\Diamond \neg (c_1 \lor c_2)$. This is not a dilemma as it stands; it is not logically impossible that both of her children are saved. But if we assume that all accessible worlds in the context are worlds realizable by Sophie, we also have to assume that $\Diamond (c_1 \lor c_2)$, because she cannot avoid sacrificing at least one child, and then we are of course contradicting (D).

If we just delete (D) from our system, we get into the minimal Kripkean system (K), where a dilemma of the form $\Diamond (a) \land \Diamond (\neg a)$ does not lead to inconsistency at the propositional level but to normative trivialization (also known as deontic explosion), i.e. that we can derive $\Diamond (b)$ for any $b$, which, on the semantic level, corresponds to the lack of accessible (permissible) worlds.

An assessment of the possibility of deontic dilemmas would require a metaethical discussion, which is beyond the scope of this thesis. The system CDM1, as I define it in chapter 2, is like SDL in that it does not allow for dilemmas. However, in chapter 6, I will discuss some ways of handling moral dilemmas that have been proposed in the literature, and I will also show how CDM1 can be changed in order to handle most moral dilemmas.

1.5.6 Can SDL Include all Deontic Concepts?

A proposition can be regarded as obligatory, permitted or forbidden, and, as we have seen, these three categories are definable in terms of each other in SDL. But are there not other normative categories into which we can classify propositions, which may be related to the traditional categories of obligation and so on, but may not be definable in terms of these? If so, should not an adequate system of logic for the normative concepts be able to handle these as well? For example, many philosophers think of the supererogatory as a normative category that forms a proper subclass of the permitted-but-not-obligatory: a supererogatory state of affairs is one that is regarded as especially good from a moral point of view but not obligatory, for example that someone makes a
great personal sacrifice in order to achieve some greater good for others. If supererogation exists, it seems that it belongs among the deontic concepts, because it is logically related to the uncontroversial deontic concepts – if a state is supererogatory, it is permitted but not obligatory – and it seems to be a concept of moral worth, e.g. a person who has realized a supererogatory state by some action is regarded as morally praiseworthy, at least if some conditions related to the person’s intentions behind the action are satisfied. If we believe that some states are supererogatory, we should also believe in states that are permitted but not optimal, and states that are genuinely morally indifferent, which, if we assume the existence of supererogation, cannot be reduced to permitted-but-not-obligatory.

As in the case of moral dilemmas, assessing the possibility of supererogation and other deontic categories beyond the traditional triad would require taking a stand on controversial metaethical issues. Moreover, the properties of the traditional deontic categories would not change if we did allow for supererogation, and neither would the relations between them, because the supererogatory is a subclass of the permitted, and because its mere inclusion into the deontic categories cannot affect e.g. the logical relations between obligation and permission. The system CDM1 will not, as I define it, include any deontic categories beyond the traditional triad. In chapter 6, I will discuss some proposed treatments of supererogation in relation to deontic logic, and I will also show how CDM1 can be extended in order to handle supererogation.

1.6 The Need of a Quantified Deontic Logic

There are some issues about quantified modal logics that remain controversial. These concern the interpretation of the individual domain, and, for systems with an identity predicate, the necessity of identity statements. In the simplest semantics for quantified modal logic, it is assumed that we have a fixed individual domain, where all possible worlds are assigned the same set of individuals. Prima facie, this seems to conflict with the view that there could have been other individuals than there actually are. It seems also to be possible to find statements where the existence of individuals is normatively evaluated, like “That individual should not have been brought into existence!” But the fixed domain interpretation has many defenders who, for example, define an extra existence predicate, which may or may not be satisfied by an individual in a world.

It is also commonly assumed in quantified modal logic that statements about the identity of individuals are necessarily true or necessarily false. But it seems that identity statements involving definite descriptions are often contingently true or false, and such statements may also be subject to deontic evaluation, like “The person sitting there should not be the prime minister of Sweden!” However, if we assume a theory of definite descriptions where those are not
seen as genuine individual terms, this type of example is no counterexample against the necessity of identity statements.

In both cases, it seems that the correct solution to these problems in deontic logic is dependent on the correct solution in modal logic in general, which is obviously beyond the scope of this thesis. In view of this, I will not investigate further into the formal properties of quantification in deontic contexts. The view that quantification in deontic contexts behaves the same way as quantification in general modal contexts has been contested by some, e.g. Castañeda. However, I will argue in chapter 5 that the arguments given by these authors fail.

Even though we obviously quantify in deontic contexts, I will present CDM1 as a system of propositional logic. I do not need quantification in order to solve the specific problems I treat, and defining a first-order logic, and then merely stipulating its quantificational properties (e.g. fixed identity and necessary identity statements) just seems like an unnecessary complication.

1.7 The Meta-Ethical Neutrality of Deontic Logic

In SDL and other systems of deontic logic, we make inferences from statements that say that something is, for example, obligatory. Usually, logical deduction is thought to be truth-preserving: if the premises are true, so is the conclusion. Does that hold for deontic logic? According to some common meta-ethical theories, normative statements cannot be true or false, because they are merely expressions of attitudes. Is deontic logic incompatible with such claims, in the sense that they exclude normative inferences? We may roughly define the following incompatibility claim:

(I) If expressivism is true, there are no valid normative inferences.

One might hold that the existence of normative inferences is so obvious that if (I) is correct, then expressivism cannot be correct. The possibility of logical relations between normative statements under expressivism has been the subject of much debate in recent times, and addressing these problems would go beyond the scope of this thesis. Many philosophers hold “minimalist” theories of truth, according to which one’s ascription of truth to a statement just signals one’s acceptance of the statement, and expressivism seems to be compatible with holding that normative statements indeed have truth-values in this sense. For well-known, relatively recent attempts to explain logical relations between normative statements under the assumption of some non-cognitivist theory, see Simon Blackburn in [6], and Alan Gibbard in [28].

If we hold that normative statements are not truth-bearers, and that logical deduction is truth-preserving, so that there, in a strict sense, can be no
logical relations between normative statements, we may give some alternative interpretation of deontic logic. For example, in the paper where SDL was first formulated, Bengt Hansson suggested a “descriptive” interpretation of the system, viz. that the formulae of SDL “describe what is obligatory, forbidden and permitted respectively, according to some undetermined system of norms or moral or legal theory” [31, p. 123].

The idea that deontic logic is about descriptive statements that stand in a certain relation to genuine normative statements, has a long tradition in the philosophical discussion. Ingemar Hedenius, who claimed to defend both expressivism and (I), tried to explain what he claims to be the illusion of inferences between normative statements by distinguishing between genuine and non-genuine imperative and normative statements [34, p. 121]. If someone says, “Do everything you have promised”, I can know from this that I have been told to do everything I have promised, and if I also know that I have promised to go to a seminar, I can conclude from this that I have been told to go to the seminar, and this can make me grasp the imperative “Go to the seminar”, which, according to Hedenius, may create the illusion of a real deductive inference between the genuine imperatives. When I grasp a genuine imperative statement, I also know that a corresponding non-genuine imperative statement is true, which is just the statement that the norm or imperative has been given, i.e. that I have been told to perform the act specified in the imperative. Normative statements are, in this aspect, analogous to imperative statements, as Hedenius claims [34, p. 128f]. Hedenius also holds that non-genuine normative statements may be expressed with the same words as genuine statements [34, p. 127]. I will not go further into the problems surrounding this type of interpretation of deontic logic.

Deontic logic, as traditionally conceived, has been challenged in relation to other meta-ethical views than expressivism. Torbjörn Tännsjö claims that deontic properties, like rightness and wrongness, are objective qualities possessed by concrete actions, and that the ordinary principles of deontic logic, e.g. the consequence principle, make no sense, given this view. He writes that “the fact that \(a_1 \cap a_3\) is obligatory does not imply logically that also \(a_1\) is obligatory (no more than the fact that a certain chair is comfortable to sit on implies logically that one of its legs is also comfortable to sit on)” [57, p. 12].

Some type of approach using a notion of “secondary” or “non-genuine” normative statements may also be put in relation to the challenge to deontic logic presented by Tännsjö’s concretism. Tännsjö has not shown that his view is incompatible with claiming that we also may use normative language to reason about necessary and sufficient conditions for the performance of those actions that are the primary right-bearers. One might argue that we must have some linguistic resource of this type, in order to be able to use language to get people to perform right actions. For example, we might use an expression like “It ought to be that \(a\)” in order to state that \(a\) is a necessary condition for
the performance of a right action, and we might then claim that deontic logic consists in the study of the logical properties of these concepts.

1.8 Philosophical Applications of Deontic Logic

It is not easy to find moral philosophers who actually make use of deontic logic, e.g. in the formulation of different ethical theories. The lack of consensus in deontic logic in comparison with ordinary modal logic, and, of course, ordinary first-order logic, has prevented its use as a tool. As a moral philosopher, one can hardly know which, if any, interpretation of $\Box(a)$ it is that captures what one wants to say, and one cannot suppose one’s audience to understand formulae with such expressions without detailed explanation.

1.8.1 Understanding Hume’s Law with or without Deontic Logic?

In his book *The Is-Ought Problem*, Gerhard Schurz has used deontic logic to explicate and defend different versions of Hume’s Law, which roughly says that logical derivations of normative conclusions from only factual premises are impossible. This principle has been challenged by A.N. Prior, who has pointed to valid arguments like

1. Tea drinking is common in England.
2. Therefore, tea drinking is common in England or all New Zealanders ought to be shot.

This argument is of the structure $a \vdash a \lor \Box(b)$. We may think of the conclusion, $a \lor \Box(b)$ as non-normative, but then we have to say that the argument $\neg a, a \lor \Box(b) \vdash \Box(b)$ is a valid argument with purely non-normative premises and a conclusion that cannot be said to be non-normative. A common response to this type of argument has been to say that Hume’s Law should be interpreted as claiming that no consistent set of purely descriptive premises logically implies a purely normative conclusion. $a \lor \Box(b)$ would be an example of a mixed sentence. Schurz holds that it is unsatisfactory to consider only pure sentences, because even mixed sentences play an important part in moral reasoning. He posits his “General Hume Thesis” [53, p. 78], which says that any possibly mixed conclusion from a purely descriptive premise set is “completely $\Box$-irrelevant”, which roughly means that the variables of the conclusion that can be arbitrarily substituted in the $\Box$-scope *salva validitate* (e.g., $a \lor a \lor \Box(b)$ would still be valid if we substituted $\neg b$ for $b$ in the conclusion).

Schurz also discusses another “irrelevance” principle, that of *I*(s)-O*(ught*)-triviality. Proposition $a$ is said to be an Ought-trivial conclusion of the premise set $D$, if $a^{\neg \Box}$, viz. the result of omitting all deontic operators in $a$, is also derivable from $D$ [53, p. 132]. An obvious example of a trivial derivation is the
derivation of $\Box T$, by the principle (RN). Perhaps a more philosophically interesting example is that if we, in Schurz’s combined modal and deontic logic, assume (OC) $\Box(a) \supset \Diamond(a)$, which might be seen as an interpretation of “ought implies can”, we can from $\neg\Diamond(a)$ derive $\neg\Box (a)$. This is an ought-trivial conclusion, because we can also derive $\neg a$. Schurz shows that, in deontic predicate logic with identity and fixed individual domain, a version of the General Hume Thesis is violated, but that these violations are trivial [53, p. 188].

Some writers have argued that Hume’s Law may be analyzed without the use of deontic logic. An example is Charles R. Pigden, who holds that the logical autonomy of ethics “is merely an instance of the more general thesis that logic is conservative” [45, p. 129], and thinks this is sufficient for analyzing the formal aspects of Hume’s Law. An occurrence of an expression $\phi$ in the conclusion of a valid inference, $K \vdash X$, is, according to his definition, vacuous iff we can uniformly substitute for $\phi$ any expression $\psi$ of the same grammatical type to get a new sentence $X'$, such that $K \vdash X'$ is also a valid inference. After giving this definition, Pigden proves, for first-order predicate logic, that “[a] predicate or propositional variable cannot occur non-vacuously in the conclusion of a valid inference unless it appears among the premises” [45, p. 136].

Håkan Salwén, in his dissertation Hume’s Law, discusses different versions of that law without using deontic logic, and adopts Pigden’s principle as a “formal” version of the law. He mentions Schurz’s work, but says that “neither Schurz nor other deontic logicians have given any philosophical motivation for considering the logical machinery that in Schurz’s case, becomes extremely complex” [51, p. 139].

How could we capture a notion like Schurz’s I(s)-O(ught)-triviality if we did not have some resource for operating on the contents of obligations? This notion is intended to cope with prima facie counterexamples against Hume’s Law that come from principles such as (RN) or (D) in deontic logic. Pigden is skeptical about deontic logic. He finds the principle (RN) absurd, and argues that (D) cannot be logically true either, because contradictory obligations can arise in some moralities [45, p. 139]. He concludes, after briefly considering some alternatives to SDL, that there “is no such thing as deontic logic; no principles of inference peculiar to, and pervasive of, all moral or normative reasoning” [45, p. 141]. For my part, I think many defenders of metaethical theories like prescriptivism would claim that something like (D) must be true, because of the action-guiding nature of morality (and try to explain apparent dilemmas in terms of conflicts between prima facie principles), and would also be interested in an interpretation of Hume’s Law that, in a non-ad hoc way, can avoid conflict with that claim. Schurz’s interpretations of Hume’s Law by means of deontic logic may be interesting in the context of analyzing the implications of such, not obviously absurd, metaethical views.

Schurz’s irrelevance notion captures some inferences that cannot be handled in Pigden’s simpler machinery. We can derive an instance of a truth in deontic logic, like the (RE) instance $\Box(a \lor b) \leftrightarrow \Box(b \lor a)$, from any premises, but
it will not be replaceable in Pigden’s sense. (RE) is much less controversial than (RN) or (D). If we think that deontic logic captures something essential of normative reasoning, in the sense that there are some valid inferences that can only be captured by means of some kind of deontic logic, we have good reason to use deontic logic in the analysis of Hume’s Law.

1.9 Some Adequacy Conditions

Given the aims stated in section 1.3, and given that we need a special deontic logic at all (as I have argued in 1.4 that we do), we can define some adequacy conditions for a system of deontic logic. In chapter 3, I will argue that an adequate deontic logic should contain indexed operators of moral requirement, indexed to sets of possibilities, such that some kind of rule of regularity hold. Let us call this requirement ROI (Regular Operators with Index). Moreover, I will argue that it should also contain possibility-indexed operators for which the rule of regularity does not hold. Call this requirement NROI. For simplicity and keeping with philosophical tradition, it is also desirable if we can handle the SDL problems with a system that fulfills the following conditions.

C The system is built on classic logic.

P The deontic operators operate on propositions.

MC The system uses the material conditional for conditional obligations.

I leave it open whether the following conditions also should be satisfied.

SE The system handles supererogation.

Dil The system handles genuine obligation dilemmas, where we have two jointly unsatisfiable states that are both obligatory.

In the end, in section 7.3, I will compare how the different systems discussed in the dissertation fare in these aspects.
2. The System CDM1

2.1 Introduction

In this chapter, I will give an account of the system CDM1, which I propose as a way to meet the aims stated in the introductory chapter. I will first give an informal presentation of the central features of CDM1 with an example formalization of an ethical argument of some complexity, and then give a formal specification of the system.

2.2 Informal Specification

2.2.1 General Features of CDM1

I propose a deontic system where a state is obligatory just in case its realization is both necessary for moral optimality and guaranteed to be all-things-considered morally better than its non-realization, in a certain sense which will be explained below. I will use a kind of so-called neighborhood semantics, of the type described by Krister Segerberg in [54], for the deontic operators. Obligation is indexed to a point with a set of possible states described by normal modal operators. In many contexts, it will be appropriate to think of the set of possible states as the set of propositions compatible with the possible courses of action open to an agent at a given time, but the system itself does not entail any such specific interpretation of the alternative sets. The system allows for arbitrarily many indexed modal and deontic operators, in order to handle the fact that different alternative sets may give rise to different sets of obligations, which, as I shall explain, is important e.g. in the treatment of the Chisholm Paradox. Formally, the indices are natural numbers. To improve readability, I will often denote indices in examples with variables.

I assume that the necessity operators, \( \square_i \), have the properties of the modal operator in the modal system (T), i.e. the minimal Kripkean modal logic (K), extended with the axiom schema \( \square(a) \supseteq a \), which says that necessity implies truth. On the semantic level, this corresponds to the condition of reflexivity of the accessibility relation, viz. that the world from which the necessity operator is evaluated is itself a possible world.

This yields a connection between the different indices: for any indices \( i, j \), \( \square_i(a) \) is inconsistent with \( \square_j(\neg a) \). What is necessary relative to one index, is at least possible relative to all indices.
I take the operators of obligation, $\Box_i(a)$, as primitive. For these operators, the consequence principle is invalid. We only assume (ORE), replacement of necessary equivalents, the aggregation schema (OAND), which says that $\Box_i(a) \land \Box_i(b)$ implies $\Box_i(a \land b)$, and a corresponding schema for disjunction, (OOR), which says that $\Box_i(a) \land \Box_i(b)$ implies $\Box_i(a \lor b)$.

I also define operators of moral requirement, $\text{MR}_i(a)$, that state necessary conditions for fulfilling obligations. A proposition $a$ is $i$-required iff there is some $i$-obligatory $b$, such that $\Box_i(b \supset a)$. Thus, $\Box_i(a) \supset \text{MR}_i(a)$ (but not the converse) is a theorem for any $i$. To handle the relation between the $\Box$- and MR-operators, we need quantification over propositions in the object language, and I will use an adaption of Kit Fine’s extension of the standard modal systems in [24] for that purpose. I assume the schema (DPos), saying that whatever is $i$-required is $i$-possible.

Intuitively, the obligation operator should be interpreted so that a proposition $a$ is $i$-obligatory iff $a$ represents a state that is not only morally required, i.e. necessary in order to obtain a morally optimal outcome, but also sufficiently strong to $i$-necessarily ensure an all-things-considered moral improvement of the outcome, compared with $\neg a$, if as little else as possible is changed, within the limits of what is $i$-possible. In the following, I will argue that this interpretation of obligation is able to handle the SDL problems, in accordance with the aims described in 1.3.

From the point of view of the classical Utilitarian, $\Box_i(a)$ may be interpreted so that $a$ is necessary to attain the highest possible net balance of pleasure over pain in the whole world, and sufficient to bring with it a higher net balance of pleasure over pain in the whole world than we would have obtained if $a$ had been false, regardless of whatever else is the case, within the limits of what is $i$-compatible with $a$ and $\neg a$. If we have the worlds $w_1 \models a \land b, w_2 \models a \land \neg b, w_3 \models \neg a \land b, w_4 \models \neg a \land \neg b$, accessible in relation to $w_1$, with the hedonic ordering $w_1 > w_2 > w_3 > w_4$, $a$ and $b$ will both come out as obligatory at $w_1$, if we assume that nothing else than the difference in truth-value of the propositions $a$ and $b$ affects the degree of similarity between these worlds, i.e. $w_3$ is a minimally different $\neg a$-world and $w_2$ a minimally different $\neg b$-world in relation to $w_1$, and $w_4$ a minimally different $\neg a$-world in relation to $w_2$, and a minimally different $\neg b$-world in relation to $w_3$. On the other hand, if we have the ordering $w_1 > w_2 > w_4 > w_3$, $a$ but not $b$ will come out as obligatory, because $b$ gives a better outcome than $\neg b$ would have given only if $a$ is realized.

If we have a deontological morality, with a list of duties, $\Box_i(a)$ may be interpreted so that $a$ is necessary to attain complete fulfillment of duties, and always realizes more duties than $\neg a$ would have realized, without coming into conflict with other duties of equal or higher moral importance. A crucial feature of many deontological moralities is their agent-relativity: what is

---

1Thanks to Johan Gustafsson for the proposal of including the disjunctive schema.
important to the moral status of e.g. my actions, is what duties I fulfill or violate by my actions, and not the consequences of my actions for the general duty-fulfillment in the world. We may then, of course, say that only the duties realized by the actions of a specific agent are relevant to what is obligatory in relation to an index. Note that CDM1 is not suited to handle deontological moralities where duties may come into conflict so that we face genuine dilemmas; these problems are discussed in chapter 6.

This interpretation will give us the properties of the \( \circ \)-operator in CDM1, which is what we need to handle the SDL problems in accordance with the aims stated in section 1.3. Closedness under strict equivalence, i.e. the (ORE) schema, should be unproblematic. (DPos), (OAND) and (OOR) should also be valid. The conjunction of (DPos) and (OAND) captures the intuitions of moral requirement lying behind CDM1, i.e. that individual requirements, as well as sets of requirements, should be possible to fulfill. If \( a \) and \( b \) both are \( i \)-necessary to obtain a morally optimal outcome, they should both be jointly \( i \)-satisfiable. And if \( a \) and \( b \) both are moral requirements, that are \( i \)-sufficient to make an all-things-considered improvement over \( \neg a \) and \( \neg b \), and \( a \land b \) is possible, \( a \land b \), as well as \( a \lor b \) should also be sufficient for making an improvement over \( \neg(a \land b) \), or \( \neg(a \lor b) \), because both \( a \land b \) and \( a \lor b \) must be made true by at least one of \( a \) and \( b \) being made true. On the other hand, closedness under strict or logical consequence does not hold. For example, \( a \lor b \) is a logical consequence of \( a \), but even if \( a \) necessarily would make the world better, the realization of \( a \lor b \) by \( b \) need not make the world better. The Alf Ross Paradox, whose treatment in relation to CDM1 is discussed in section 4.2.1, may be seen as a demonstration of this feature of normative reasoning in ordinary language.

One way to think about the \( \circ \)-operator might be to say that a \( \circ_i(a) \) holds just in case we have some \( a \)-world better than all \( \neg a \)-world, and any \( a \)-world is morally better than any minimally different \( \neg a \)-world. One might then wish for a representation of the \( \circ \)-operators in a formal semantics defined directly in terms of some value ranking of worlds and a similarity relation between worlds. However, finding such a semantics that validates the (OOR) property in an intuitive way has proven hard. Instead, I represent the \( \circ \)-operator by means of neighborhood models, where the property of being a requirement minimally sufficient to ensure an all-things-considered improvement is taken as a primitive property of propositions, i.e. sets of worlds.

The main idea with the distinction between the obligation and the requirement operators is that the truth of \( \circ_i(a) \) must not only be necessary but also sufficient for something of positive moral value to be realized, which is not the case with \( \text{MR}_i(a) \).

Permission and prohibition are defined relative to moral requirement, as they are usually defined relative to obligation in deontic logic. A proposition, \( a \), is permitted relative to context \( i \) (written \( \text{P}_i(a) \)) iff its negation is not morally required, and \( a \) is forbidden (written \( \text{F}_i(a) \)) iff \( \neg a \) is required.
The idea that permission should be defined in terms of strict compatibility with any obligatory proposition has, to my knowledge, never been worked out in any formal system, but was suggested by Sven Danielsson in [20, p. 6].

When the operators of permission and prohibition are defined in relation to the MR-operator in this way, one might perhaps get the impression that the MR-operator really is the central obligation operator, and that the ○-operator is introduced as some kind of *ad hoc* device, in order to avoid the SDL problems. However, we could have omitted the MR-operator and instead have given equivalent definitions of the P- and F-operators directly in terms of the ○-operator.

\[(P)\ P_i(a) \leftrightarrow_{\text{def}} \neg \exists b(\circ_i(b) \land \Box_i(b \supset \neg a))\]

\[(F)\ F_i(a) \leftrightarrow_{\text{def}} \exists b(\circ_i(b) \land \Box_i(b \supset \neg a))\]

The (DPos) axiom could also be stated in a simple way in terms of the ○-operator.

\[(DPos)\ \circ_i(a) \supset \Box_i(a)\]

At the same time, I believe MR-operator captures some natural concept of requirement or duty, which may be given as a reason for not omitting it. For example, I believe that necessity verbs like “must” in natural English often fill a function similar to the MR-operator, as in statements like: “You must φ in order to do your duty, because you are obliged to the conjunction of φ and ψ.”

Many, but not all, of the relations between obligation, permission and prohibition in SDL are preserved for the ○-operator in relation to the P- and F-operators in CDM1. It should be obvious that we have one side of the equivalences commonly used to permission and prohibition in SDL preserved in CDM1. \(\circ_i(a) \supset \neg P_i(\neg a)\) and \(\circ_i(a) \supset F_i(\neg a)\) are both valid, though their converses are not. I will further discuss the advantages of this approach in 4.2.1.

My distinction between obligation and requirement operators bears some similarity to Erik Carlson’s distinction between different consequentialist principles in his book *Consequentialism Reconsidered* [7]. According to his primary KO-principle of obligation and his KR-principle of rightness, an action ought to be done or is right only if its outcome is “optimal no matter what other actions the agent performs, in the same situation” [7, p. 103]. But he also defines a PO-principle of obligation, according to which an action ought to be done iff it is, in a certain sense necessary in order to achieve an optimal outcome, and a PR-principle of rightness, according to which an action is right iff it is compatible with achieving an optimal outcome [7, pp. 35-38].
These principles are defined in relation to a consequentialist morality with a framework of agency and alternatives, and they are not intended to give a characterization of normative reasoning in general. However, the PO- and PR-principles stand in a relation to moral optimality rather similar to the MR- and P-operators in CDM1. If \( a \land MR_i(\neg a) \) (or, equivalently, \( a \land \neg P_i(a) \)) holds, we know that some \( i \)-obligation is not satisfied. Similarly, in Carlson’s framework, if action \( \phi \) is performed, and PO \( \sim \phi \) (where \( \sim \phi \) means the non-instantiation of \( \phi \) [7, p. 6]), or \( \neg PR \phi \) holds, an optimal outcome has not been achieved. An important difference between my \( \bigcirc \)-operators and Carlson’s KO-principle is that if KO \( \phi \) holds, the performance of \( \phi \) guarantees an optimal outcome, relative to the situation, while the truth of \( a \land \bigcirc_i(a) \) is fully compatible with some other \( i \)-obligation being dissatisfied. We could define an operator of “moral sufficiency” that would play a role similar to Carlson’s KR, and an operator of “sufficient requirement” similar to Carlson’s KO:

\[
\text{(MS)} \quad MS_i(a) \leftrightarrow \text{def} \ P_i(a) \land \forall b(\bigcirc_i(b) \supset \Box_i(a \supset b))
\]

\[
\text{(MSR)} \quad MSR_i(a) \leftrightarrow \text{def} \ MS_i(a) \land MR_i(a)
\]

MS\(_i\)(a) holds just in case a is permitted, and it is \( i \)-necessary that if a is true, any \( i \)-obligation is satisfied, and the truth of a thus, in a natural sense, guarantees moral optimality. MSR\(_i\)(a) means that a is both necessary and sufficient for moral optimality.

However, I think these would be too strong to capture any concept of rightness or obligation common in ordinary language. When we say a thing like “You may stay in bed”, or “You ought to stay in bed”, we normally do not want to exclude that there are things you are forbidden to do that are compatible with your staying in bed, such as sending insulting messages from your phone (cf. [20, p. 5]).

Any instance of the following deontic schema, which states that whatever is morally required is permitted, and whose (non-indexed) counterpart is assumed as an axiom schema in standard deontic logic, follows from the assumptions above.

\[
\text{(D) For any } i : \vdash MR_i(a) \supset P_i(a)
\]

To have MR\(_i\)(a) \( \land \) MR\(_i\)(\( \neg a \)), would mean that we had some propositions, b and c, such that \( \bigcirc_i(b) \land \bigcirc_i(c) \land \Box_i(b \supset a) \land \Box_i(c \supset \neg a) \), i.e. the obligations to b and c would not be jointly satisfiable. Because aggregation is valid for the obligation operators, we could then derive \( \bigcirc_i(b \land c) \), and MR\(_i\)(a \( \land \) \( \neg a \)), which would contradict (DPos).

Common theorems in deontic logic that are invalid for the \( \bigcirc_i \)-operators include the following:

\[
\text{Common theorems in deontic logic that are invalid for the } \bigcirc_i \text{-operators include the following:}
\]

\footnote{These characterizations are rather rough, because I do not use Carlson’s specific terminology, which is unnecessary for the present purposes.}
1. $\vdash \Box_i(a) \supset \Box_i(a \lor b)$
2. $\vdash \Box_i(a \land b) \supset \Box_i(a)$
3. $\vdash [\Box_i(a \supset b) \land \Box_i(a)] \supset \Box_i(b)$

However, all of these become valid if we substitute MR for $\Box$ in the consequent of the main implication. $\vdash \Box_i(\top)$ is not valid. MR$_i(\top)$ is not valid either, because CDM1 does not exclude the possibility that there are no obligations for some $i$, i.e. that there is no $b$ such that $\Box_i(b)$, which implies that there is no $a$ such that MR$_i(a)$. But any instance of $\vdash \Box_i(a) \supset \text{MR}_i(\top)$ is valid.

Any instance of the following holds.

($P\top$) For any $i$: $P_i(\top)$

Some philosophers, like J.L. Mackie in [42], have proposed so-called error theories of moral judgment. These theories say, roughly, that moral judgments are false or lack truth-value, because they ascribe to things objective properties that do not exist. CDM1 is intended to be neutral between different meta-ethical positions, as far as possible. It might be thought that ($P\top$) conflicts with error theory, because we have, at least, to acknowledge that some things are permitted.

However, if $\neg\text{MR}_i(\top)$ holds, we can say that any proposition $a$ is trivially $i$-permitted. CDM1 does not come into conflict with an error theory stating that all moral judgments, except for judgments stating trivial permissions, are false, or that trivially permissive judgments are not really moral judgments at all, in any sense relevant to the error theory.$^3$

We should remember that $P_i(a)$ just means that $a$ does not conflict with any moral obligation. We sometimes issue statements of the form “You are permitted to $\phi$” in order to express stronger concepts, implying that there are certain obligations for others not to interfere with your $\phi$-ing, and so on. I do not intend to capture any such concepts with the P-operators.

Any instance of the following holds.

($P\text{Pos}$) For any $i$: $\vdash \Box_i(a) \supset [P_i(b) \supset \Diamond_i(b)]$

For any $b$ such that $\Box_i(\neg b)$ holds, $\Box_i(a \supset \neg b)$ also holds for any $i$-obligatory $a$, which, by definition of the MR- and P-operators, means that MR$_i(\neg b)$, i.e. $\neg P_i(b)$, holds.

$\vdash [\Box_i(a \supset b) \land \Box_i(a)] \supset \Box_i(b)$ is valid. We will use this schema in the treatment of the so-called contrary-to-duty paradoxes.

When an alternative context, $i$, is given, we can define the following translation $t$ of natural English deontic statements into the language of CDM1, where

---

$^3$Some error theorists claim that moral judgments have no truth values at all, not even in a minimal sense, and this problem is of course not solved by distinguishing between trivial and non-trivial permissions: see section 1.7 for further discussions about the meta-ethical neutrality of deontic logic.
$p$ is a sentence in natural English (if we assume that $t$ already has been defined for non-deontic statements, in the style of [3, p. 35]):

1. $t(\text{It ought to be that } p) = \Box_i(t(p))$
2. $t(\text{It is obligatory that } p) = \Box_i(t(p))$
3. $t(\text{It is permitted that } p) = P_i(t(p))$
4. $t(\text{It is forbidden that } p) = F_i(t(p))$

It is assumed that statements of the form “$X$ ought to perform action $\phi$” can be paraphrased as “It ought to be that $X$ sees to it that $X$ performs $\phi$”. See section 1.5.4 for a discussion about alleged problems with this approach. In examples, I will often use simplified paraphrases where this has no bearing on the argument. If we want to investigate the formal properties of this seeing-to-it relation, we can add an operator in the style of e.g. Stig Kanger in [39] or John Horty in [35]. I discuss issues related to this in section 2.2.2 below, but I do not think this is relevant to my main arguments about the SDL problems.

The MS- and MSR-operators defined above may, as I said, not answer to any natural concept of permission or obligation. But we can use the MSR-operator to illustrate a property of CDM1.

Remember that the idea behind the obligation operators was that $\Box_i(a)$ means that $a$ is both morally required and ensures the existence of something of positive moral value. This is also the case if MSR$_i(a)$. We should expect the following to hold:

**(OMSR)** For any $i$: $\vdash \text{MSR}_i(a) \supset \Box_i(a)$

**(OMSR)** does indeed hold in CDM1, which can be proved the following way:

1. $\text{MR}_i(a) \land \text{MS}_i(a)$ By definition of MSR.
2. $\exists b \Box_i(b) \land \Box_i(b \supset a)$ 1, by definition of MR.
3. $\forall b \Box_i(b) \supset \Box_i(a \supset b)$ 1, by definition of MS.
4. $\Box_i(b) \land \Box_i(b \supset a)$ 2, by rules of propositional quantification.
5. $\Box_i(b)$ 4, by propositional logic
6. $\Box_i(b \supset a)$ 4, by propositional logic
7. $\Box_i(a \leftrightarrow b)$ 3, 5, by rules of propositional quantification
8. $\Box_i(a \leftrightarrow b)$ 6, 7, by modal logic
9. $\Box_i(a)$ 5, 8, by ORE

It also holds that any two propositions $a$ and $b$, such that MSR$_i(a) \land \text{MSR}_i(b)$ holds are strictly $i$-equivalent i.e. $\Box_i(a \leftrightarrow b)$ holds. MS$_i(a) \land \text{MR}_i(b)$ entails $\Box_i(a \supset b)$.

### 2.2.2 The Interpretation of Necessity

We may say that Smith had a duty to welcome the new students yesterday, but not today, and that Jones never has had any duty to welcome new students, because the welcoming is not an alternative relative to Smith today, and was
never relative to Jones at any time. We may also say that Smith had a duty to be present at the lecture yesterday if she promised to, but not that she automatically had any duty to be present at the lecture. But if she now in fact has promised, we would say that she now really has a duty to be present, because she has promised. This can be explained by claiming that when we speak of something being a duty, we presuppose a set of alternatives, and hold something a duty only if it would occur if some member of that alternative set were realized.

CDM1 is intended to represent such a concept of obligation, where a proposition is obligatory only if it is possible relative to a set of possible states of affairs. When holding something possible, or impossible, one is warranted to make certain inferences. For example, you cannot hold that Smith is obliged to \( a \), and that Smith therefore is obliged to \( b \), because \( \neg a \) is impossible for Smith, and at the same time claim that Smith is obliged to \( \neg a \), unless you change alternative sets. In a speech context, such a change may, as in the examples above, be announced by a change of subject or time-reference in the main clause in a normative sentence.

One might ask for a more detailed analysis of what this possibility consists in. A full investigation into the conditions for the morally relevant possibility of something would require more than one thesis itself. One example of a theory that would have to be assessed in order to accomplish this could be a strong incompatibilism, according to which we are never obliged to do anything other than we actually do, if the proposition (Det), a definition of determinism adapted from Peter van Inwagen [60, p. 186], below is true, because we can never under such circumstances do anything other than we actually do.
For every state of world, there is a proposition that expresses the state of the world at that instant, and if \(a\) and \(b\) are any propositions that express the state of the world at some instants, then the conjunction of \(a\) with the laws of physics entails \(b\).

Given such an incompatibilistic theory, \((\text{Det})\) would imply \(a \leftrightarrow \square_i(a)\) for any proposition \(a\) and any index \(i\) referring to the alternatives of an agent.

\(\square_i(a)\) should not be interpreted as saying that \(a\) is true and that some agent or group of agents is unable to ensure that \(\neg a\). Then it would be similar to a common interpretation of Peter van Inwagen’s N-operator, where \(N(a)\) means that \(a\) is true and nobody ever has had a choice about \(a\). If we interpret “having a choice about \(a\)” as being able to ensure the falsity of \(a\), it can be convincingly shown that e.g. the principle \(N(a) \land N(b) \supset N(a \land b)\), which is valid for any Kripkean necessity operator, is invalid for the N-operator. (For a discussion about the logical properties of the N-operator, see Erik Carlson’s article [8].)

We might think that facts about the abilities of agents to ensure certain states may be relevant in assessing the possibilities relative to one context. But in such cases, we should use some agent-relative seeing-to-it operator, in the style of e.g. Stig Kanger in [39] or John Horty in [35], in the scope of the necessity operator. The statement that agent \(x\) sees to it that \(a\) is written \(\text{Do}(x, a)\) with Kanger’s operator. Let \(h\) and \(t\) stand for a pair of mutually exclusive random events (let us say, \(h\) : the coin lands heads, and \(t\) : the coin lands tails), such that agent \(x\) is able to ensure \(h \lor t\) but not able to ensure \(h\) and not able to ensure \(t\). We could interpret this as \(\square_i(\neg \text{Do}(x, h) \land \neg \text{Do}(x, t)) \land \Diamond_i(\text{Do}(x, h \lor t)))\), which would be fully consistent in CDM1 with an added Do-operator. The same holds if we use the cstit-operator employed by John Horty in his deontic systems in [35], where \(x \text{cstit} : a\) is read as that \(x\) sees to it that \(a\): \(\square_i(\neg x \text{cstit} : h \land \neg x \text{cstit} : t) \land \Diamond_i(x \text{cstit} : h \lor t))\) would be fully consistent.

In this thesis, I only assume a minimal theory of necessity/possibility that I think will suffice for dealing with the deontic paradoxes, such as the contrary-to-duty paradoxes, where the distinction between different alternative sets is relevant. I think any reasonable conception of necessity in this context should at least satisfy the properties of the necessity operator in Kripke’s system (T), e.g. that necessity is closed under logical consequence and that it implies truth.

As mentioned above, we have thus a minimal connection between the indices, such that \(\forall a[\square_i(a) \supset \Diamond_j(a)]\) holds for any indices. We can assume further connections and say that one index \(j\) is included in another index \(i\) if it holds that \(\forall a[\square_i(a) \supset \square_j(a)]\). This may be useful e.g. if \(j\) and \(i\) refer to the possibilities open at different time-points, where \(j\) comes later than \(i\), and we think that the possibilities monotonically shrink as time passes. However, connections of this type should not be assumed as logical truths in a system of deontic logic suitable for the purposes stated in 1.3, for reasons I will explain in 4.4.1.
2.2.3 Conditional Obligations in General

Assume that we have a common natural language argument of the form:

1. It is obligatory that $b$, if $a$.
2. $a$
3. Therefore, it is obligatory that $b$.

From $\Box_i(a)$ and $\bigcirc_i(a \supset b)$, we can derive $\bigcirc_i(b)$. If it is $i$-obligatory that $b$ is the case, if $a$ is the case, $b$ is $i$-obligatory in the case when $a$ is something that is $i$-settled. The normal interpretation of arguments like the one above is intended to be $\bigcirc_i(a \supset b), \Box_i(a) \vdash \bigcirc_i(b)$. In many cases, we might also interpret the conditional premise with narrow scope for the obligation operator, like $a \supset \bigcirc_i(b)$. Then it suffices, of course, to interpret the second premise as $a$ in order to derive the conclusion. However, narrow scope does not work in some cases, such as the Chisholm situations discussed in 4.4.1.

2.3 Example of a Philosophical Argument in CDM1

In order to give an idea about the relation between CDM1 and argumentation of the kind pursued in moral philosophy, I will here give an CDM1 representation of a well-known type of argument from this latter area. The aim of this type of argument is to show that a certain moral theory is false because it can, given certain assumptions, be used to derive certain normative conclusions that conflict with other statements that we also assume true under these circumstances. In my example, I will interpret the argument as showing that a certain version of welfare-maximizing consequentialism is false, because it implies that there is a requirement to realize a state of affairs where we have a high total sum of welfare in the world, but with a very large number of existing people and a very low average level of welfare, even though it would be possible to realize a state with a moderate number of existing people and a high average level of welfare. This is, of course, an application of Derek Parfit’s well-known challenges for population ethics in Reasons and Persons [44, p.387–90].

Let $u$ denote the consequentialist principle the argument is aimed to refute, $iOpt$ denote a proposition stating that an optimal outcome, in terms of the total sum of welfare, is realized, $z$ denote a proposition stating that the universe contains a high total welfare sum and a low average welfare, and $a$ denote a proposition stating that the universe contains a lower total welfare sum but a higher average welfare (because of the smaller number of people living at a higher level of welfare) than $z$.

1. $u \supset \forall p \bigcirc_i(p) \leftrightarrow \Box_i(p \leftrightarrow_i Opt)$
2. $\Box_i(z \leftrightarrow_i Opt), P$
3. $\Box_i(z \supset \neg a), P$
4. $\bigcirc_i(a), P$
5. \( F_i(z), 3, 4, \text{DPos} \)
6. \([u, P]\)
7. \(|\bigcirc_i(z), 1, 2, \text{PC}\)
8. \([\bot, 5, 7, \text{DPos}\]
9. \(\therefore \neg u, 6–8, \text{RAA}\)

The principle \( u \) implies that proposition \( z \) is \( i \)-obligatory, and the argument then shows how the assumption that \( a \), which is not \( i \)-compatible with \( z \), is obligatory can be used as a reductio of the principle \( u \). The premise about the incompatibility of \( z \) with \( a \) is assumed to lie inside the scope of a necessity operator, so that we can use it in arguments with the (DPos) schema.

### 2.4 Formal Specification

#### 2.4.1 Language

The CDM1 language \( \mathcal{L} \) contains the following primitive symbols:

1. Propositional variables: \( a, b, c \ldots \)
2. The propositional all-quantifier: \( \forall \)
3. The connectives of propositional logic: \( \neg, \supset \)
4. Delimiting symbols: ()
5. Standard necessity operators: \( \Box_i, \Box_j, \Box_k \ldots \) (with indices \( i, j \in \mathbb{N} \))
6. Operators of deontic necessity: \( \text{MR}_i, \text{MR}_j, \text{MR}_k \ldots \)
7. Operators of obligation: \( \bigcirc_i, \bigcirc_j, \bigcirc_k \ldots \)

We may then define the set of formulae of \( \mathcal{L} \) as the smallest set \( \Phi \) of the above symbols such that:

1. If \( a \) is a propositional variable, \( a \in \Phi \).
2. If \( a \in \Phi \), \( \neg a \in \Phi \).
3. If \( a \in \Phi \) and \( b \in \Phi \), \( a \supset b \in \Phi \).
4. If \( a \in \Phi \), \( \Box_i a \in \Phi \), for any \( i \).
5. If \( a \in \Phi \), \( \bigcirc_i a \in \Phi \), for any \( i \).
6. If \( p \) is a propositional variable and \( a \in \Phi \), \( \forall b(a) \in \Phi \).

We may further define the propositional connectives \( \land, \lor, \leftrightarrow \) and the existential quantifier as usual in classical logic, and the possibility, requirement, obligation, permissibility and prohibition operators:

1. \( \Diamond_i a \leftrightarrow \text{def} \neg \Box_i \neg a \), for any \( i \).
2. \( \text{MR}_i a \leftrightarrow \text{def} \exists b(\bigcirc_i(b) \land \Box_i(b \supset a)) \), for any \( i \).
3. \( P_i a \leftrightarrow \text{def} \neg \text{MR}_i \neg a \), for any \( i \).
4. \( F_i a \leftrightarrow \text{def} \text{MR}_i \neg a \), for any \( i \).
2.4.2 Deductive System

We assume the following axioms and derivation rules for CDM1, in addition to all rules and axioms in a classical system of propositional logic.

The Kripkean axiom schema for each standard necessity operator.

(K) For any $i$: $\vdash \Box_i(a \supset b) \supset (\Box_i(a) \supset \Box_i(b))$

Necessitation rule for the standard necessity operators.

(RN) For any $i$:

\[
\begin{array}{c}
\vdash a \\
\hline
\vdash \Box_i(a)
\end{array}
\]

The truth axiom for the standard necessity operators.

(T) For any $i$: $\vdash \Box_i(a) \supset a$

Replacement of necessary equivalents for the obligation operators.

(ORE) $\vdash [\Box_i(a) \land \Box_i(a \leftrightarrow b)] \supset \Box_i(b)$

Aggregation schema for the obligation operators.

(OAND) For any $i$: $\vdash \Box_i(a) \land \Box_i(b) \supset \Box_i(a \land b)$

Schema for the introduction of obligatory disjunctions.

(OOR) For any $i$: $\vdash \Box_i(a) \land \Box_i(b) \supset \Box_i(a \lor b)$

The axiom of deontic possibility.

(DPos) For any $i$: $\vdash MR_i(a) \supset \Diamond_i(a)$

Axiom schemata and the generalization rule for the propositional quantifier, adapted from Fine:

(EV) $\vdash \forall p(a) \supset a[q/p]$, where $q$ is free for $p$ in $a$.

(KV) $\vdash \forall p(a \supset b) \supset (\forall p(a) \supset \forall p(b))$

(NV) $\vdash a \supset \forall p(a)$, where $p$ is not free in $a$.

(BF) For any $i$: $\vdash \forall p \Box_i(a) \supset \Box_i \forall p(a)$

(G)

\[
\begin{array}{c}
\vdash a \\
\hline
\vdash \forall p(a)
\end{array}
\]

(DBF) In addition to Fine’s proof system, there is also the Barcan Formula for the obligation operators: For any $i$: $\vdash \forall p \Box_i(a) \supset \Box_i \forall p(a)$
2.4.3 Semantics

We assume the following semantics, CDM1S, for CDM1. A CDM1S model is an ordered quintuple \(< W,R,S,P,V >\), where \(W\) is a set of worlds, and \(R\) and \(S\) are sets of binary relations, \(R_i,R_j,R_k\) and \(S_i,S_j,S_k\) (with the indices \(i,j,k\in\mathbb{N}\) such that for any \(R_i,R_j\subseteq W\times W\), and for any \(S_i,S_j\subseteq W\times\mathcal{P}W\).

\(P\) is the set of propositions: a set of non-empty subsets of \(W\).

So, the \(R\)-relations correspond to the usual accessibility relation in a Kripkean modal logic. The \(S\)-relations are used to single out the set of propositions that, relative to an index, are \(i\)-desirable in the sense that they are both necessary for the realization of a morally optimal state and sufficient to guarantee that their realization means an all-things-considered moral improvement of the world, compared with the realization of their negations, as described in section 2.2.1.

We assume the following conditions for the \(R\)- and \(S\)-relations.

1. For any pair of relations \(< R_i,S_i >\), any world \(w\), and any non-empty set of worlds \(X\subseteq P\), if \(wS_jX\), there is some world \(w'\), such that \(wR_iw'\) and \(w'\in X\). Intuitively, this means that for each \(i\)-desirable state, there should be an \(i\)-possible way to realize that state.

2. For any \(S\)-relation \(S_i\), any world \(w\), and any sets of worlds \(X\) and \(Y\), if \(wS_jX\) and \(wS_jY\), then \(wS_jX\cap Y\). The intersection of two desirable states also stands for something desirable.

3. For any \(S\)-relation \(S_i\), any world \(w\), and any sets of worlds \(X\) and \(Y\), if \(wS_jX\) and \(wS_jY\), then \(wS_jX\cup Y\). The union of two desirable states also stands for something desirable.

4. For any pair of relations \(< R_i,S_i >\), any world \(w\) and any non-empty set of worlds \(X\subseteq P\), if \(wS_iX\), for any non-empty set of worlds \(Y\), such that \(X\cap\{w'|wR_iw'\} = Y\cap\{w'|wR_iw'\}\), \(wS_jY\) holds. For any two states, such that one \(i\)-necessarily is realized iff the other is, one denotes something that is \(i\)-necessary and \(i\)-sufficient for the realization of something morally \(i\)-desirable iff the other does.

5. The \(R\)-relations are reflexive.

\(V\) is a function assigning exactly one value from the set \(\{0,1\}\) to formulae in \(\Phi\) relative to worlds \(w\in W\), according to the following conditions. We say that proposition \(a\) is true at world \(w\) just in case \(V(a,w) = 1\), and false at world \(w\) just in case \(V(a,w) = 0\).

1. If \(p\) is a propositional variable, \(V(p,w) = 1\) or \(V(p,w) = 0\).
2. For any formula \(a\), \(V(\neg a,w) = 1\) if \(V(a,w) = 0\); otherwise \(V(\neg a,w) = 0\).
3. For any formulae \(a\) and \(b\), \(V(a\lor b,w) = 1\) if \(V(a,w) = 0\) or \(V(b,w) = 1\); otherwise \(V(a\lor b,w) = 0\).
4. For any formula \(a\) and any \(\Box\)-operator \(\Box_i\), \(V(\Box_i(a),w) = 1\) if \(V(a,w') = 1\) for every \(w'\) such that \(wR_iw'\); otherwise \(V(\Box_i(a),w) = 0\).
5. For any formula \(a\) and any \(\bigcirc\)-operator \(\bigcirc_i\), \(V(\bigcirc_i(a),w) = 1\) if \(wS_i\{a\}\), where \(|a| = \{w' \in W|V(a,w') = 1\}\); otherwise \(V(\bigcirc_i(a),w) = 0\).
6. For any propositional variable \( p \), and any formula \( a \), \( V(\forall p a, w) = 1 \) if \( V'(a, w = 1) \), in any other model \( < W, R, S, P, V' > \), where the valuation \( V' \) is like \( V \) except for the interpretation of \( p \); otherwise \( V(\forall p a, w) = 0 \).

Thus, for the \( \Box_i \)-operators, we use a standard Kripkean semantics, where the accessibility relation holds between worlds, while we for the non-normal \( \bigcirc_i \)-operators use a “neighborhood semantics”, where the relation holds between worlds and propositions, i.e. sets of worlds.

We assume that \( P \) is, to adapt Fine’s terminology, closed under formulae, viz. that for any formula \( a \in \Phi \), \( |a| = \{ w \in W | V(a, w) = 1 \} \in P \). The reason why we do not assume that \( P \) just is the power set of \( W \) is that CDM1S then would be impossible to axiomatize. Fine has shown that we, if we make this assumption in systems with propositional quantifiers based on a weak modal system like (T), can use the modal accessibility relation and the propositional quantifier to translate second-order arithmetic into the system, and then apply the old incompleteness results [24, p. 343f].

2.4.4 Soundness and Completeness

The \( \Box_i \)-operators, and the propositional quantifier are based on Kit Fine’s system (T\( \pi \)). Fine has shown that this system is sound and complete with respect to Kripkean semantics for the standard system (T), extended with a set of propositions closed under formulae, and the given truth-conditions for formulae with quantifiers.

The (ORE) schema is valid in CDM1S. Two propositions \( a \) and \( b \) are necessarily equivalent relative to \( i \) just in case the intersection between the proposition \( |a| \) and the \( i \)-accessible worlds is the same as the intersection between \( |b| \) and the \( i \)-accessible worlds. We have assumed that if \( wS_i |a| \) holds in such a case, corresponding to the condition that \( a \) is obligatory, also \( wS_i |b| \) holds.

The (OAND) schema is also valid in CDM1S, because \( |a| \cap |b| = |a \land b| \) for any propositions, \( a \) and \( b \).

The (OOR) schema is also valid in CDM1S, because \( |a| \cup |b| = |a \lor b| \) for any propositions, \( a \) and \( b \).

An instance of (DPos), like MR\( _i (a) \supset \Diamond_i (a) \), could only be falsified in a case where MR\( _i (a) \) but not \( \Diamond_i (a) \) holds. It is easy to see that this is impossible in any CDM1S model, given the truth conditions of MR\( _i (a) \) and \( \Diamond_i (a) \) and given that \( wS_i |a| \) implies that there is some \( |a| \)-world \( w' \), such that \( wR_iiw' \).

Hence, we conclude that CDM1 is sound with respect to CDM1S, i.e. that \( \models_{CDM1} a \) implies \( \models_{CDM1S} a \) for any formula \( a \in \Phi \).

We can prove the completeness of CDM1 with respect to CDM1S the usual way for modal logics, by defining a canonical model for CDM1. The canonical CDM1 model C is a model where the set of worlds is the set of all maximal sets of CDM1-consistent formulae. If C is an CDM1S model, any formula \( a \) that is a consequence of a set of formulae \( \Gamma \) in every CDM1S model is thus derivable from \( \Gamma \) in CDM1. (For a detailed description of this method of
proving the completeness of a modal system, consult some textbook in modal logic.)

If we have the canonical CDM1 model, C, where (DPos) is valid, there cannot be any pair of relations in C, \(< R_i, S_i >\) and any non-empty subset \(|a| \in P\), such that \(wS_i|a|\) but also \(|a| \cap \{w' \in W | wR_i w'\} = \emptyset\) holds (which would imply that C is not an CDM1S model). Then we would have some \(a\) such that \(MR_i(a)\) would be true but \(\Diamond_i(a)\) would be false, which contradicts (DPos).

If (ORE) is valid in C, there cannot be any pair of relations in C, \(< R_i, S_i >\), and any pair of propositions \(|a|\) and \(|b|\) such that \(|a| \cap \{w' \in W | wR_i w'\} = |b| \cap \{w' \in W | wR_i w'\}\) and \(wS_i|a|\) hold, but \(wS_i|b|\) does not hold. Then, \(\bigcirc_i(a) \land \Box_i(a \leftrightarrow b)\) would be true, but \(\bigcirc_i(b)\) false, which contradicts (ORE).

For the axiom schema (OAND), it is well-known that we can use canonical models to prove that the schema, in conjunction with the rule of replacement of logical equivalents, which follows trivially from (ORE), is complete with respect to the class of all neighborhood frames closed under intersection. The same holds for the schema (OOR) and neighborhood frames closed under union.

Hence we conclude that a canonical CDM1 model is a CDM1S model, and that CDM1 thus is strongly complete with respect to CDM1S, i.e. that \(\Gamma \models_{CDM1S} a\) implies \(\Gamma \vdash_{CDM1} a\) for any formula \(a \in \Phi\), and any set of formulae \(\Gamma\) with each member in \(\Phi\).
3. CDM1 vs. Some Other Systems

3.1 Introduction

This chapter is to serve as a prelude to the next chapter, where I discuss how CDM1 handles the features of normative reasoning that were presented in the introduction as problematic for SDL, given my stated aim to provide a formal representation of normative reasoning in natural language. In this context, I will compare CDM1 to some other deontic systems that are similar to CDM1 in some respects, and have served as inspiration for it, but dissimilar in other respects, namely Hector-Neri Castañeda’s logic and Lewis-style dyadic systems of deontic logic. In this chapter, I will make a rather informal presentation of the central properties of these systems. Both types of systems contain some resources absent in both SDL and CDM1: Castañeda’s logic contains non-propositional entities, so-called practitions, and the Lewis-style systems contain dyadic deontic operators with a special non-material conditional. In the next chapter, I will argue that these extra resources do not give the mentioned systems any significant advantage over CDM1 in treating the SDL problems. My intention with the comparison is by no means to give any sort of comprehensive survey of all possible solutions to these problems in deontic logic; however, I think the comparison will make the distinguishing features of CDM1 clearer.

3.2 Hector-Neri Castañeda’s Deontic Logic

We say that we are obliged to certain things under certain circumstances. We may, for example, be obliged to give money to some organization if people are starving. It may be natural to say that the former italicized clause expresses something like an imperative, a practical statement, while the latter describes a circumstance. Hector-Neri Castañeda held that the distinction between acts practically considered and acts considered as circumstances is of crucial importance for deontic logic.

Castañeda writes that he uses the term practition to “refer generically to both prescriptions and intentions” [9, p. 43], but we have to explain what he means by those terms. A prescription is something that can form the practical content of an imperative. It is “a structure of an agent and an action connected in . . . a demanding way” [9, p. 40]. One might obtain “[p]ure expressions of prescriptions” by making an infinitive clause of an imperative, so the prescrip-
tion corresponding to “Mary, go home” is “Mary to go home” [9, p. 40]. Cas-
ánteda states that when he speaks about intentions, he speaks about the content
of the operation that also can be expressed by the word “intention” [9, p. 40].
This operation is analogous to a prescription in that it presents a demand of
an agent, but it differs in that it is first-person directed. If Smith intends to kill
Jones, Smith, who would express it by saying “I shall kill Jones”, has an in-
tention of the form “I to kill Jones”. So, practitions are something that present
demands of agents and form the content of imperatives and intentions, and,
according to Castañeda, also the content of normative sentences. The other
fundamental entity, which Castañeda contrasts with practitions, is the propo-
sition. This term, I think, largely has its common philosophical meaning. A
typical proposition would be “Mary goes home”. In Castañeda’s deontic logic,
only deontic statements with practitions inside the scope of the deontic opera-
tor are considered well-formed. This is a fundamental difference between his
logic and standard deontic logic, where the obligation operator operates on
propositions, as explained in the introductory chapter.

A practition may contain propositions as well, and not only, for example, a
simple infinitive clause. If B and A are well-formed practitions and p a proposi-
tion, ¬B, p ∧ B, B ∧ p, B ∧ A and ∀B are well-formed practitions too [10, p.
75]. Propositions may be moved in and out of the scope of the deontic op-
erator, so that ∪i(p ⊃ A) ↔ (p ⊃ ∪i(A)), where ∪i is Castañeda’s indexed
obligation operator, as explained below, is valid.

An action in a practition may be either practically considered, and analyzed
as a perhaps smaller practition, or considered as a circumstance, and analyzed
as a proposition or an identifier. An action A is practically considered in a
practition X iff “A appears in X in the scope of the prescriptive copula” [9,
p. 209]. On the other hand, a is considered as a circumstance in X iff a occurs
in X but is not practically considered in X and does not appear “as a part of
reference to, or characterization of, some entity, or entities” [9, p. 210]. Cas-
anteda gives the following example of an action, which appears in the same
practition first practically considered, and then considered as a circumstance.
“Peter, do the following: if it rains, close the windows, and if you-close-the-
windows uncover the skylight.” [9, p. 209]. We may well see “if it rains, close
the windows, and if you-close-the-windows uncover the skylight” as one big
practition. On the other hand, “close the windows” in “if you close the win-
dows”, cannot be seen as a practition, because it does not have a logically
imperative form where it appears in this sentence. Castañeda calls a practition
like the one above, which is made up of both practitions and propositions,
“mixed”. The latter condition in the definition of when an action is considered
as a circumstance is exemplified by the prescription “Peter, return to Smith the
book you stole from him” [9, p. 209]. This prescription contains a reference to
Peter’s stealing of a book, but this action is only considered as an “identifier”,
because it enters “as constitutes of the reference to a certain book” [9, p. 210].
Only actions can be practically considered in practitions, but propositions in

40
general may appear as circumstances or identifiers in the same way as actions do.

A practition may be a part of a deontic proposition, by which Castañeda seems to mean simply any proposition of the form “It is obligatory/forbidden/permitted/optional that A”, with the deontic property in question specified as legal, moral, institutional etc. [9, p. 190] 1. Castañeda holds that a deontic proposition is made up of a practition combined with the deontic predicate itself.

The concept of a practition may still seem unclear. What does Castañeda mean by “demanding way”, for example? Perhaps, we can say that to accept a demand is to act on it, if one is able, which is different from accepting a proposition. If someone says to me “You will go to the seminar tomorrow”, and I say “Yes, I will” and then do not go, this shows that my belief was wrong, but if someone says “Go to the seminar tomorrow”, and I say “Yes”, but yet do not go, this shows that I must have changed my mind. If we ask what the imperative “Go to the seminar” has in common with the norm “You ought to go to the seminar”, the answer might be that they share the same practical content, which can be accepted or rejected by performing or not performing a certain act. This does not necessarily imply some strong motivational internalism regarding norms. One might argue that e.g. a proposition stating a moral norm may be accepted as true, even by someone who does not accept the practical content of the norm.

Castañeda develops a quantified deontic logic built on his practition/proposition-distinction [9, ch. 9]. Castañeda has different indexed deontic languages and systems for different normative systems (e.g. legal and moral oughts), and each system is denoted $D_i^{**}$, whose non-quantified fragment is $D_i^c^{**}$. In this section, it will be appropriate to focus on those non-quantified systems. For any $D_i^{**}$-system, there is a model consisting of a triple $\langle W_0, W, I \rangle$, where $W$ is a set of “possible deontic worlds”, with the actual world $W_0$ as a member, and $I$ is a function that assigns truth-values to propositions and values of “legitimacy” to practitions. Castañeda does not believe in iterated deontic operators, and the truth-value of a deontic proposition is always evaluated relative to the actual world, $W_0$. $\circ_i A$ is true at $W_0$ iff the practition $A$ is legitimate at every other world in $W$. That a practition is legitimate at a world does not mean that it is obeyed in that world; rather, it means that it holds as a sort of active demand at that world. Accepting a practition $A$ as actually legitimate is like accepting an imperative to $A$. It is the legitimacy values of practitions that differ across the worlds in $W$; any (non-deontic) proposition has the same truth-value at every world2. This enables propositions to be moved in and out of the scope of deontic operators, as Castañeda wants. We

---

1 I do not think that he holds that this specification always has to be explicit in a sentence expressing a deontic proposition.

2 This does not seem to commit Castañeda to deny inter-world supervenience for norms. Even if we think of the different possible deontic worlds as existing entities, it is different practi-
may think of W as the set of all combinations of action demands, or imperatives, whose satisfaction are compatible with fulfilling the requirements of the normative system i. This will perhaps be clearer if we consider Castañeda’s conception of an “unqualified” ought, denoted \( \Box_i A \), which is true at \( W_0 \) iff \( A \) is legitimate at every world in W, including \( W_0 \). Castañeda’s idea is that we may combine different \( D_i \star \star \), which may come in conflict with each other: if \( \Box_i A \wedge \Box_j \neg A \wedge \Box_1 A \), then we ought, all things considered to, \( A \). Belzer and Loewer offer a somewhat more complicated model to formally represent such combinations of systems [5, p. 391–94].

3.3 Lewis-style Dyadic Systems

In deontic systems with dyadic obligation operators, obligation is not absolute but relative to some given circumstances. The obligation operator is of the form \( \Box(a/b) \), read “\( a \) is obligatory given \( b \)” (where \( a \) and \( b \) are propositions), rather than \( \Box(a) \), read “\( a \) is obligatory simpliciter”, as in e.g. SDL. This type of system was first introduced by von Wright in 1956 [62].

David Lewis has provided a semantics for this type of system in terms of possible worlds [41, p. 96–104]. He uses an adaption of his semantics for counterfactual conditionals (for which he also defines a corresponding family of “V-logics”, with “V” standing for “variable strictness” of the conditional [41, p. 118–136]) to deal with dyadic deontic operators. In this type of model, a possible world is assigned a set of “spheres”, i.e. sets of possible worlds nested into each other, and closed under union and intersection. For evaluation of deontic statements, the worlds are ranked in terms of comparative betterness, with smaller and smaller spheres containing better and better worlds. A statement like \( \Box(a/b) \) is true at a world \( i \) if there is some sphere \( S \) assigned to \( i \) such that \( S \) contains at least one world where \( b \) is true and the material conditional \( b \supset a \) is true at every world in \( S \). These truth conditions are the same as for counterfactual conditionals in Lewis’s model. The difference is that, in the counterfactual case, we use a ranking of comparative similarity instead of betterness. Some assumptions about the spheres that are made for counterfactuals, such as “centering”, i.e. that the smallest sphere around any world \( i \) contains \( i \) itself, do not hold in the deontic case; we cannot plausibly hold that any world (ours, for example) is among the best possible world relative to itself. This is analogous to denying reflexivity for the accessibility relation in Kripkean models for standard deontic logic. Lewis also defines truth conditions for unconditional obligation in terms of his semantics. \( \Box(a) \) is true at \( i \) iff \( a \) is true at every world in some non-empty sphere assigned to \( i \).

\[ \text{tions and not different norms that hold in different worlds that are identical in their descriptive aspects.} \]
Lewis-style semantics is often characterized as claiming that a conditional obligation statement \( \Box(a/b) \) is true at \( i \) if \( a \) is true at all the \( b \)-worlds that are best relative to \( i \). This may give a better intuitive understanding of the ideas behind the semantics than the truth-conditions given above, but it is also somewhat simplified. From these truth-conditions and the properties of spheres, it follows that if \( b \supset a \) holds at all worlds in one \( i \)-sphere containing \( b \)-worlds, it also holds in all smaller spheres (containing better worlds, relative to \( i \)). However, Lewis denies what he calls the “Limit Assumption” for deontic conditionals (as he does for counterfactual conditionals), viz. the assumption that there always exists a best \( b \)-world, or a set of the best \( b \)-worlds, relative to one world \( i \). Instead, Lewis thinks, we may have an infinite ascent of better and better worlds.

Different writers on deontic logic have made elaborations of Lewis-style semantics. Fred Feldman’s MO-operator, presented in *Doing the Best We Can*, is indexed to agent and time. The unconditional obligation, \( \text{MO}_{st}(a) \), means that there is some world \( w' \) such that \( w' \) is “accessible” to agent \( s \) at time \( t \), \( a \) is true at \( w' \), and that there is no world \( w'' \) accessible to \( s \) at \( t \) where \( a \) is not true and \( w'' \) is at least as good as \( w' \) [21, p. 38]. The conditional obligation \( \text{MO}_{st}(a/b) \) means that there is some \( st \)-accessible world where \( a \wedge b \) is true, and no at-least-as-good \( st \)-accessible world where \( \neg a \wedge b \) is true [21, p. 87].

Another example of a deontic system modeled on Lewis-style semantics is Marvin Belzer’s and Barry Loewer’s system 3-D, whose language “adds to propositional logic a dyadic deontic operator \( \Box_i(q/p) \) and a temporally indexed monadic operator \( !_{i[t]}(p) \)” [5, p. 400], where the index \( i \) stands for different normative systems and \( t \) for points of time. They introduce a ranking of possible worlds (relative to the different normative system), and their idea is that \( \Box_i(q/p) \) is true iff \( q \) is true in all the best accessible worlds where \( P \) is true. Belzer’s and Loewer’s !-operator is about what is required *tout court*. The rule

\[
\textbf{(Key)} \quad \Box_i(q/p), S_i(p), U_{it}(q, p) \vdash !_{it}(q)
\]

is valid in 3-D. \( S_i(p) \) says that \( p \) is “settled”, i.e. that it is unalterable at \( t \), and \( U_{it}(q, p) \) says that the conditional obligation is not defeated by any further facts: there is no \( r \) such that \( S_i(r) \) and \( \neg \Box_i(q/p \wedge r) \). A statement of “absolute” obligation like \( \Box_i(q) \) is an abbreviation of \( \Box_i(q/\top) \).
4. The SDL Problems in the Different Systems

4.1 Introduction

In this chapter, I will discuss how the SDL problems presented in the introduction are handled in CDM1 and the other systems discussed in the previous chapter. I will argue that SDL handles these problems better than any of the other systems.

4.2 The Alf Ross Paradox

4.2.1 The Alf Ross Paradox in CDM1

For the \(\square\)-operators, the derivation behind the so-called Alf Ross Paradox, which would take us from \(\square_i(a)\) to \(\square_i(a \lor b)\), is invalid, because the consequence principle does not hold for the obligation operator. In the special case when \(\square_i(\neg b)\) holds, we can, however, derive \(\square_i(a \lor b)\) from \(\square_i(a)\). Under any reasonable assumptions about necessity, “You ought to post the letter” will entail “You ought to post the letter or give it to a married bachelor”. The disjunctive obligation could then only be satisfied by the first disjunct.

For the MR-operators, the corresponding derivation is unrestrictedly valid, i.e. the schema MR\(_i\)(a) ⊃ MR\(_i\)(a ∨ b) is valid. Thus, \(\square_i(a)\) also entails MR\(_i\)(a ∨ b). If \(a\) is a necessary condition for fulfilling the \(i\)-obligations, so is \(a \lor b\). As we have defined prohibition and permission, if \(a\) is obligatory, \(\neg(a \lor b)\) is forbidden and not permitted, and this seems intuitively right. That \(a\) is forbidden does not entail that we are under an obligation to \(\neg a\), but it entails that \(a\) is morally required. The Alf Ross Paradox may be seen as an illustration of the necessity of making the distinction between moral obligations, for which some consequence principle does not hold, and mere moral requirements, for which such a principle does hold, as in CDM1. We can also derive P\(_i\)(a ∨ b) from P\(_i\)(a). It may seem odd to derive “It is permitted to post the letter or burn it” from “It is permitted to post the letter”, but I think an explanation in terms of pragmatic speech rules works better for this so-called paradox of free choice permission than for the original Alf Ross Paradox. I will discuss this issue in 4.2.3.
4.2.2 The Consensus about the Alf Ross Paradox

Castañeda writes about the Alf Ross Paradox in [10, p. 64ff], and expresses opinions fairly typical of deontic logicians. In his system, a non-propositional counterpart of the \((\lor I_d)\) theorem that lies behind the Alf Ross Paradox is validated by his version of the general consequence principle, where \(A\) and \(B\) are practices:

\[(Dr11.1)\] If \((A \supset B)\) is a theorem, so is \((\lozenge_i(A) \supset \lozenge_i(B))\)

Castañeda says that this has no paradoxical consequences at all, and that the false impression that it may come from a “erroneous form of semantical atomism” – people tend to only look at the inferred conclusion \(\lozenge_i(A \lor B)\) and forget that the agent is still bound by the premise \(\lozenge_i(A)\).

Most deontic logicians have agreed with Castañeda’s verdict that the Ross Paradox is no serious matter. It is indeed odd to assert that it is obligatory to \(A\) or \(B\) if we hold that it is obligatory to \(A\), but Castañeda, and others, have claimed that this can be explained with an appeal to pragmatic reasons. According to pragmatic rules of communication, we should not withhold relevant information, which we do when we assert that it is obligatory to \(A\) or \(B\) while withholding that it is obligatory to \(A\), just like when we assert an indicative statement like “\(a\) or \(b\)” while believing in \(a\).

4.2.3 Danielsson’s Arguments Against the Ross Inference

The pragmatic explanation of the Ross Paradox may seem prima facie plausible. However, Sven Danielsson has argued that if you, when you are obliged to do \(a\), do \(b\), which is incompatible with \(a\), or else forbidden, then you, if \((\lor I_d)\) is valid, still do something you ought to do, which he finds unacceptable [17, p.47].

The prospects for finding a way out of the Ross Paradox by giving sentences like “\(a\) or \(b\) is obligatory” more complex translations into the SDL language seem bad. Giving a more complex translation of the disjunctions within obligation operators like \(\lozenge_i(a \lor b)\) and \(\neg(\lozenge_i(a) \lor \lozenge_i(b))\) does not change the fact that some obligation is satisfied if \(a \lor b\) is true.

The Ross inference is not unconditionally valid if obligation is only applied to open alternatives, as with Nuel Belnap’s and John Horty’s operator dstit (deliberative seeing-to-it) in [36]. Statements about what an agent \(\alpha\) ought to do are interpreted as \(\lozenge[a \text{dstit} : a]\), which means that \(\alpha\) ought to see to it that \(a\). \(\lozenge[a \text{dstit} : a]\) is true iff \([a \text{dstit} : a]\) holds in every optimal possible world-history, relative to an index, and \([a \text{dstit} : a]\) means that \(\alpha\) chooses to bring about \(a\) so that \(\neg a\) has to be open to \(\alpha\). Now, you do not necessarily choose \(a \lor b\) if you choose \(a\), because \(a \lor b\) may be something already settled, if \(b\) is settled. \(\lozenge[a \text{dstit} : a \lor b]\) thus cannot be derived from \(\lozenge[a \text{dstit} : a]\), even though Horty and Belnap accept closedness under logical consequence for the
the consequence of the satisfaction of obligation, which we do not want to have as an obligation, is something that is not settled; you have a choice about the disjunction to post the letter or burn it, and then it holds that if you choose to post, you also choose to post or burn. In such cases, this type of solution does not avoid the alleged paradox.

In a later paper, Danielsson introduces a restricted consequence principle [19, p. 24]:

\[(A6) \quad [\bigcirc_i(a \land b) \land M_i(b \land \neg a) \land M_i(\neg b \land \neg a)] \supset \bigcirc_i(a)\]

\(M_i(a)\) means that \(a\) is an alternative relative to the index \(i\). To claim that \(a\) is an alternative, in Danielsson’s sense, is incompatible with claiming either \(a\) or \(\neg a\). This notion of alternatives will be discussed further in connection with the contrary-to-duty paradoxes, in 4.4.4.

With A6, we cannot derive \(\bigcirc_i(a \lor b)\) from \(\bigcirc_i[(a \land (a \lor b))]\) (which of course is equivalent with \(\bigcirc_i(a)\)), because we cannot realize \(a \land \neg(a \lor b)\), or from \(\bigcirc_i[(a \land b) \land \neg b]\), because we cannot realize \(\neg(\neg b) \land \neg(a \lor b)\). Danielsson wants to block \((\lor Id)\), while preserving \((\land Id)\bigcirc_i(a \land b) \supset \bigcirc_i(a)\) when the conjuncts can be realized and not realized independently of each other, so that, for example “you ought to buy milk” follows from “you ought to buy milk and bread”.

Danielsson says that he wants to capture the intuition that “for instance, you close the door seems to be a proper part of you close the door and turn on the light in a sense in which you close the door or turn on the light is not a proper part of you close the door and either close the door or turn on the light” [19, p. 25]. However, Jörg Hansen has shown that this solution cannot avoid the Ross Paradox [30, p. 225f]. If we assume another obligation \(\bigcirc_i(c)\), which is independent of \(\bigcirc_i(a)\) and can be fulfilled, and not fulfilled, even if \(a\) and \(b\) are not fulfilled, we can still derive \(\bigcirc_i(a \lor b)\), via the derivation of \(\bigcirc_i[((a \lor b) \land (a \lor \neg b)) \land c]\), where the first conjunct is equivalent with \(a\), thus yielding \(\bigcirc_i[(a \lor b) \land ((a \lor \neg b) \land c)]\), by permutation of the conjuncts. In Hansen’s example, \(c\) stands for being a faithful husband, while \(a\) and \(b\) stands for posting and burning the letter, as in Ross’s original example.

If we reject \(\lor Id\), we cannot keep the usual definitions of the relations between obligation, permission and prohibition. If we hold \(\bigcirc(p)\), we surely want to say that \(\neg(p \lor b)\) is something that is forbidden and not permitted. If it is obligatory that the letter is posted, it is forbidden that it is neither posted nor burnt. But with the usual definitions of the forbidden, \(F(p) \leftrightarrow_{\text{def}} \bigcirc(\neg p)\), and the permitted, \(P(p) \leftrightarrow_{\text{def}} \neg \bigcirc(\neg p)\), \(F(\neg(p \lor b))\) and \(\neg P(\neg(p \lor b))\) are both equivalent to \(\bigcirc(p \lor b)\). Sven Danielsson has then suggested that we define the operator of prohibition \(F(p)\) so that \(p\) is inconsistent with some \(q\), such that \(\bigcirc(q)\), and permission, \(P(p)\), so that \(p\) is not prohibited [20, p. 6].

---

1In his later work [35], Horty has abandoned the dstit-operator for the interpretation of ought-to-do-sentences, and instead uses the cstit-operator, which is closed under logical consequence.
In CDM1, I follow the main aspects of Danielsson’s proposal, as I have explained. One might ask why our deontic concepts have this structure, which seems to be more complicated than the relations between our ordinary modal concepts. Well, there seems to be an intuition that one cannot satisfy a duty by realizing the, from a moral point of view, worst possible state in a situation, even if one can satisfy a duty by realizing a suboptimal state, as in contrary-to-duty situations. If there is a connection between failure to fulfill duties and blameworthiness, we may think that if you fulfill a duty by realizing \( p \), you must at least be less blameworthy than you would have been if you had not realized \( p \).

We may, in this context, note the so-called paradox of free choice permission. If \( a \) is \( i \)-permitted, then it follows in CDM1 that \( a \lor b \) is too. Is it not just as odd to claim “You are permitted to go to the movies or to rob the bank” when you are permitted to go to the movies as it is to claim “You ought to post the letter or burn it” when you are obliged to post the letter? It would be odd, but in this case, a pragmatic explanation in terms of the rules of speech, or perhaps an alternative translation of disjunctive norms, maybe of the form \( P_i(a) \land P_i(b) \) for “it is permitted that \( a \) or \( b \)”, may be more appropriate than in the case of obligation. We have a similar oddity in the case of alethic modalities: \( \Diamond(a \lor b) \) follows form \( \Diamond(a) \), but it would seem nonsensical to assert “It is possible that I am in Stockholm in one hour or that I am in Babylon in ten minutes” if I think it would be possible for me to be in Stockholm, but not in Babylon, in one hour.

The reason why I do not accept pragmatic explanations or alternative translations in the obligation case is, as noted above, that I am inclined to think that an obligation cannot be satisfied by the realization of the worst alternative in a situation, even if an obligation can be satisfied by the realization of a bad alternative, as in the case of contrary-to-duty obligations. Considerations of this type do not seem relevant when it comes to permissions, because permissions are not entities that can be “satisfied” in a way that makes the agent morally praiseworthy. You act in accordance with some permission if you go to the movies or rob a bank, if you are permitted to go to the movies, but your mere acting in accordance with a permission does not imply that the you are either praiseworthy or blameworthy for this. It may be that the fact that you have satisfied a permitted disjunction entails that you are not blameworthy because you have satisfied the disjunction; but I do not find it counterintuitive to say that you, if you have robbed a bank, are not blameworthy because you have gone to the movies or robbed a bank, but only because you have robbed a bank.
4.3 The Good Samaritan Paradoxes

4.3.1 The Good Samaritan Paradoxes in CDM1

The Good Samaritan Paradoxes are based on the consequence principle. Therefore, they are not valid for the $\bigcirc_t$-operators, for which this principle is invalid.

Take an example of Åqvist’s Paradox of the Knower.

(Know) You ought to know that Frida has cancelled her concert (for which you have bought tickets).

The statement that fulfilling your obligation requires you to know that Frida has cancelled her concert, may, letting $c$ stand for the proposition that Frida has cancelled her concert and $JB_p(a)$ for the statement that $p$ has a justified belief that $a$, be interpreted as:

(KnowObl) $\bigcirc_{you}(c \land JB_{you}(c))$

We cannot derive $\bigcirc_{you}(c)$ from this, because the consequence principle is invalid in CDM1. If we want to claim that Frida should not have cancelled the concert, we can interpret this judgment as referring to another alternative set, like $\bigcirc_{Frida}(\neg c)$.

On the other hand, for the MR-operators, the traditional versions of the Good Samaritan Paradox will be valid. These operators are aimed to express what conditions are necessary for fulfilling a duty.

(KnowReq) $MR_{you}(c \land JB_{you}(c))$

From (KnowReq), which is entailed by (KnowObl), we may derive $MR_{you}(c)$. However, this only means that we cannot claim that you have an obligation to know $c$ and at the same time claim that $\neg c$ is compatible with your fulfilling your obligations\(^2\), which should not be paradoxical at all. In relation to Frida, $MR_{Frida}(\neg c)$ is trivially entailed by $\bigcirc_{Frida}(\neg c)$, so we need the different indices in order to handle this type of situation.

So-called adverbial versions of the Good Samaritan Paradox, maybe first presented by James W. Forrester in [25], have been thought impossible to handle in some systems of deontic logic that can handle the ordinary Samaritan Paradoxes. We have the following set of sentences:

(Adv1) Smith ought not to murder Jones.

(Adv2) If Smith murders Jones, Smith ought to murder Jones gently.

\(^2\)Provided, of course, that “fulfilling your obligations” is read *de re* and not *de dicto*: if $\Box_{you}(\neg c)$ were the case, you would have another set of obligations, which does not include any obligation to know $c$. 

49
(Adv3) Smith murders Jones.

We may try to formalize these as:

(Adv1.1) $\Box_s(\neg Mu(j))$

(Adv2.1) $Mu(j) \supset \Box_s(Mug(j))$.

(Adv3.1) $Mu(j)$

We can then, by modus ponens, derive $\Box_s(Mug(j))$. Something seems to have gone wrong. We do not want to say that Smith ought to murder Jones, not even gently. With (Adv 1.1), we can also derive the compound obligation $\Box_s(\neg Mu(j) \land Mug(j))$, which is impossible to fulfill, and thus conflicts with the (DPos) schema. Well, if we interpret (Adv2) literally the way we have done, it is not surprising that (Adv1)–(Adv3) turn out to be inconsistent.

Such things as (Adv2) are often uttered in contexts where we want to persuade a person like Smith, who we do not think we can persuade into accepting (Adv1), to at least kill Jones gently.

(Adv2.2) $\Box_s(Mu(j) \supset Mug(j))$.

(AdvPerm) $P_sMu(j)$

(Adv2.2) may be an adequate interpretation of an utterance like (Adv2). It would presumably come out as true, given a plausible normative theory, because any $s$-possible world where $Mu(j) \supset Mug(j)$ is true would be better than any minimally different world where $Mu(j) \supset Mug(j)$ is false, i.e. where Jones is murdered in a non-gentle way. One might ask what the point of communicating such a sentence is, if we think that murdering itself is impermissible. Well, we might imagine a situation where Smith thinks that murdering Jones is permitted, and thus accepts the following.

(AdvPerm) $P_sMu(j)$

We might think that we cannot convince Smith to accept (Adv1.1) but that we can convince Smith to accept (Adv2.2). With (AdvPerm) and (Adv2.2), Smith may derive that the killing of Jones should be done in a gentle way. We may still think that Smith really should not kill Jones at all. Note that (Adv2.2) does not follow from (Adv1.1) in CDM1, because the consequence principle is invalid.

4.3.2 The Scope Solution in SDL

The problematic statement in Good Samaritan-type examples, like “You should help Smith, who has been robbed”, is often a relative sub-clause.
One might argue that relative sub-clauses, which appear to be embedded in modal contexts, generally should be analyzed as logically not being inside the scope of the modal operator. If we do not keep this in mind, we can derive paradoxes parallel to the Good Samaritan in other modal contexts where we want a consequence principle, but where nothing like the practition/proposition-distinction can be applied. Take this example with doxastic modality: Smith believed that the stove, which was switched on, had been switched off. This entails that the stove was switched on, but if we analyze the sub-clause as being inside the scope of the belief operator, which we may call B, and assume a principle like $(a \Rightarrow b) \Rightarrow (B(a) \Rightarrow B(b))$, we can derive that Smith believes that the stove was switched on, which is contrary to what we want to say.

One might try to handle Good Samaritan-Type Paradoxes in SDL by means of so-called scope solutions, where the proposition that we do not want to recognize as something obligatory is put outside of the scope of the obligation operator, viz. “You should help Smith, who has been robbed” might be interpreted as “Smith has been robbed ∧□(You help Smith)”. This type of solution was perhaps first recognized by Anders Wedberg in [66, p. 219f].

Castañeda has written that a scope solution of this type requires that a statement such as “There is a man x such that Arthur will kill x and Arthur has a duty to bandage x”, which, with the scope solution, can be seen as an interpretation of a statement like “Arthur has a duty to bandage the man he will murder a week hence” does not imply “There is man such that Arthur has a duty to do the following: to bandage him while it is the case (or, it being that case) that Arthur will kill him” [10, p. 54].

Obviously, Castañeda thinks that the “while”-clause must be seen as being logically inside the scope of the obligation operator, but it is not clear from his text why he thinks so. If there is such an implication, it would probably hold in other modalities where the act/circumstance-distinction cannot be applied, such as “Last night I dreamed that Mrs. I, who has two grown-up daughters but whom I have never talked to, boasted to me about both of her daughters being great runners”. This seems to entail “There is a woman such that I dreamed the following: that she boasted to me about her daughters while it is case that I have never talked to her”, and it is obvious that the fact that I have never talked to Mrs. I is not a part of what I dreamed.

However, I admit that there are well-known Good Samaritan-type examples in deontic logic, which may be harder to handle by these scope distinctions, like the Paradox of the Knower, which is of the structure “Smith did a, but ought not to have done a, and Jones ought to know that Smith did a, but Jones knowing that Smith did a, implies that Smith did a”. With the consequence principle, we can derive that Smith ought to have done a. Anders Wedberg thought that we can use the scope solution even in cases such as these. “Jones ought to know that Smith did a” can, with the classical interpretation of knowledge, according to Wedberg, be interpreted as “Smith did a” ∧ “Jones ought to
have a justified belief that Smith did a” [66, p. 219f]. It seems a bit unnatural
to cut up a predicate like “know” in this way, so that one part falls outside the
scope of the deontic operator and another part falls inside, but this may not
be a decisive argument against such an interpretation. Castañeda writes that
“a duty to know is not the same as a duty to believe and have evidence” [10,
p. 55]. But in Castañeda’s system, $p \land \circ_i B$ is equivalent to $\circ_i (p \land B)$, so he
cannot say that the scope solution does not capture the logical content of the
obligation. However, even Wedberg acknowledged that cases where we want
to say that the transformation from something bad to something good is oblig-
atory, such as “The doctor ought to remove the bullet from the patient’s brain”,
cannot be handled with the scope strategy. We have to have a reference to the
bullet inside the scope of the deontic operator to capture the fact that the duty
consists in the transformation from a state where the patient has a bullet in the
brain to a state where there is no bullet there.

The difficulties with a Wedberg-style scope solution in cases such as this
may be clearer if we consider the point of such solutions.

1. You ought to give Smith, who is ill with meningococcal infection, antibi-
otics.
2. Smith is ill with meningococcal infection $\land \circ (\text{You give Smith antibiotics})$
3. $\neg \circ (\text{Smith is ill with meningococcal infection})$

The idea with (2) is to give an interpretation of (1) that is compatible with
(3). But from (2) and (3) follows, in SDL:

1. $P[\neg (\text{Smith is ill with meningococcal infection}) \land (\text{You give Smith antibi-
otics})]$

We do not want that. I do not see how the resources of SDL could be em-
ployed to handle examples such as this.

4.3.3 Castañeda’s System

Castañeda holds that ignorance of the distinction between acts considered
practically and acts considered as circumstances is the root of most of the
so-called deontic paradoxes. He notes that we can derive paradoxes of the
Good Samaritan Type if we assume a principle like the following, with the
pseudoformal notation Castañeda uses in this discussion.

(P’) If $^*X$ performs $A^*$ entails $^*Y$ performs $B^*$, then $^*X$ is obligated$_i$ to do $A^*$
entails $Y$ is obligated$_i$ to do $B$, where ’i’ stands for an adverb denoting
the type of obligation.

This is a variant of the consequence principle in deontic logic. With this
principle, we can derive “Jones is obligated to murder Smith a week hence”
from “Jones is obligated to bandage Smith, who Jones will murder a week
hence”, if we put the whole sentence after “to” inside the scope of the obliga-
tion operator, because “Jones will murder Smith a week hence” follows from
the second sentence. This is the common Good Samaritan Paradox. But in Castañeda’s system \( P' \) is invalid.

Instead, he has the following alternative principle:

\[(P*) \text{ If practition } *X \text{ to do } A^* \text{ entails practition } *Y \text{ to do } B^*, \text{ then } *\text{It is obligatory}_i \text{ that } X \text{ do } A^* \text{ entails } *\text{It is obligatory}_i \text{ that } Y \text{ do } B^*.\]

The fully formal counterpart of this principle is the following derivation rule, where \( A \) and \( B \) are practitions:

\[(\text{Dr 11.1}) \text{ If } (A \supset B) \text{ is a theorem, so is } (\bigcirc_i (A) \supset \bigcirc_i (B))\]

With this principle, the paradoxical consequence cannot be derived, just because “Jones will murder Smith a week hence” is only a proposition and not a practition. [9, p. 215].

Castañeda’s solution may work for the standard versions of the Good Samaritan Paradox, but it is not clear that it offers any great advantage compared to CDM1, which does not assume the non-propositional entities in Castañeda’s system. Castañeda also worked out even more complicated versions of his system in order to cope with the adverbial versions of the paradox [11].

4.3.4 Lewis-style Systems

Lewis’ original system is, in principle, vulnerable to paradoxes of the Good Samaritan-type, just as SDL. Feldman tries to circumvent these paradoxes with his time- and agent-relative MO-operator. If \( A \) is unalterable for agent \( s \) at \( t \), \( a \) is also an MO-duty for \( s \) at \( t \). But this is perfectly consistent with holding that \( \neg a \) was a duty for \( s \) at an earlier time, or is still a duty at \( t \) for some other agent. Belzer and Loewer have a similar approach. Feldman also has an alternative obligation operator, \( \text{MO}^* \), where \( \text{MO}^*_{st} p \) is defined as \( \text{MO}^*_{st} p \land \text{K}_{st} \neg p \), where \( \text{K}_{st} p \) means that \( p \) occurs in a world accessible to \( s \) at \( t \) [21, p. 43]. If \( p \) is a proposition about some event that has occurred before \( t \), \( \text{MO}^*_{st} p \) will be false, because \( \neg p \) is not accessible to \( j \) at \( t \). However, Feldman uses his primitive MO throughout his book.

As we have seen, we do not get into these problems with CDM1, because the consequence principle is invalid for the obligation operators. However, if \( a \) entails \( b \), \( \bigcirc_i (a) \) implies \( \text{MR}_i (b) \) in CDM1. In some contexts, we might want to distinguish between requirements that are open, i.e. may be violated, relative to some index and fixed circumstances. We may then define a derived operator like Feldman’s \( \text{MO}^* \), \( \text{MR}^*_{i} (a) \leftrightarrow_{\text{def}} \text{MR}_i (a) \land \Diamond_i (\neg a) \). Adding \( \Diamond_i (\neg a) \) as a conjunct to the operators of permission and prohibition might also give us an interpretation of these expressions that comes closer to natural language.
4.4 Contrary-to-Duty Paradoxes

4.4.1 Contrary-to-Duty Paradoxes in CDM1

There are many variants of contrary-to-duty paradoxes, where different systems of deontic logic have problems handling situations where some duties arise because some other duties are broken. We can recall Chisholm’s example:

1. Jones ought to go to visit Smith.
2. Jones ought to notify Smith about visiting if Jones is to visit Smith.
3. If Jones is not to visit Smith, Jones ought not to notify Smith.
4. Jones is not to visit Smith.

The violation of something that is a duty relative to one context, e.g. the alternatives of an agent at a given time, may be something unavoidable relative to another context. This may occur when the duty has been violated in the past, when someone else will violate the duty in the future, or even when the agent will violate the duty in the future and cannot do anything about it now. If we read into the Chisholm example that Smith not only will neglect visiting Jones, but that deciding to visit is really impossible at a time-point $t$, this would be an example of that type of situation. In all these cases, the violated duty, and any derived duties that hold only if this duty holds, do not really hold for Jones at $t$, because their performances are not open alternatives. On this reading, the premises in the Chisholm example would be inconsistent, if we assume the same index for all obligations. But we could say that the duty to visit holds for Jones at a later time-point $t_2$, when Jones may be able to make the decision to visit Smith. We may not always now have full control of our later actions.

In other situations, it seems that we want to assume that some duties arise because other duties, whose performances remain open alternatives, are in fact broken. The Gentle Murder case in the section above is a situation of this type, and also Chisholm’s original example, in one reading different from the paragraph above – viz. if we assume that visiting Smith is under Jones’s control, but that Jones simply chooses not to visit Smith. I propose that all these cases are treated the same way as the Gentle Murder case, where the requirement to not even murder gently is not defeated because the duty not to murder will be broken. If we assume that visiting Smith is still an open alternative for Jones, Jones is still really required to notify Smith, even though this visit will never take place. But we may hope that Jones will still be reasonable enough to not both notify Smith and then neglect visiting.

If we assume that Jones in fact will not realize $g$, but that this is not something unalterable at $t$, we may have the following formalization:

1. $\Diamond_{jt}(g)
2. \circledcirc_{jt}(g \supset n)
3. \circledcirc_{jt}(\neg g \supset \neg n)
4. \neg g
5. \Diamond_{jt}(g)$
From these premises, we may derive MR_{jt}(n). Jones’s duty cannot be fulfilled without notifying, and not notifying is thus forbidden. We may note that with a narrow scope interpretation, viz. g ⊃ □_{jt}(n), this derivation would be impossible.

The following is an example of a Chisholm situation where it is assumed that Jones will, at a later time, t2, neglect the duty to go to visit Smith, but that this is not under Jones’s control at the earlier time t.

1. □_{jt2}(g)
2. □_{jt2}(g ⊃ n)
3. □_{jt}(¬g)
4. □_{jt}(¬g ⊃ ¬n)

From these premises, we may derive e.g. ¬g, that Jones will not in fact go, ¬□_{jt}(g), that Jones at t is under no obligation to go, □_{jt}(¬n), that Jones at t is under an obligation to not falsely notify Smith about going, and also MR_{jt2}(n), that Jones at t2 is morally required to notify Smith. No dilemma, viz. a statement of the form □_{i}(a) ∧ □_{i}(¬a) for some i and a, may be derived.

I think CDM1 also may be used to treat what may be called synchronous contrary-to-duty situations of the kind discussed in the debate about so-called actualist and possibilist versions of consequentialist ethics. Frank Jackson and Robert Pargetter discuss a case first described by Holly Smith, where Jones is now driving through a tunnel behind a truck, and is about to change lanes, even though this is illegal and would disrupt the traffic. If Jones were not to change lanes, it would have very bad consequences to accelerate, because she would then crash into the truck. But in the actual situation, where she will change lanes, it would have better consequences to accelerate, because this would minimize the disruption of the traffic. Now, Jackson and Pargetter hold that Jones, in this situation, ought to accelerate now, but also that “what Jones ought to do now is not-accelerate-and-not-change-lanes” [37, p. 251]. These prescriptions cannot be jointly fulfilled, but Jackson and Pargetter claim that “incompatible prescriptions directed simultaneously to the same agent are acceptable provided they are out of different sets of options” [37, p. 247]. CDM1 can handle a set of moral norms of the type Jackson and Pargetter propose, because no specific assumptions about the relation between the necessity operators and agents and time-points are built into the logic.

We may assign Jones at t two alternative sets with the indices jt and jt+. Let the proposition that Jones changes lanes be denoted c and the proposition that Jones accelerates a.

1. □_{jt}(c)
2. □_{jt+}(¬c)
3. □_{jt}(a)
4. □_{jt+}(¬c ∧ ¬a)
The premises above are perfectly consistent in CDM1. The fact that Jones will change lanes is regarded as unalterable in relation to the $jt$-set, but not in relation to the $jt+$-set.

4.4.2 Castañeda’s System

Castañeda also treats the Contrary-to-Duty Paradoxes in deontic logic by using his practition/proposition-distinction. He uses Chisholm’s paradox as an example [9, p. 218ff]:

1. Jones ought to go to visit Smith.
2. Jones ought to notify Smith about visiting if Jones is to visit Smith.
3. If Jones is not to visit Smith, Jones ought not to notify Smith.
4. Jones is not to visit Smith.

Paradoxical consequences resulting from the formalization of these premises have been much discussed in connection with various deontic systems, and Castañeda shows how such consequences can arise in his system with a careless formalization of the premises. The following principle is valid in Castañeda’s system, where $A$ and $B$ are practitions:

$$(K^*) \quad (A \supset B) \supset (\Box(A) \supset \Box(B))$$

With (K*), we can derive that Jones ought to notify Smith about visiting from the first two premises above, but we can also, from the last two premises, derive that Jones ought not to notify Smith, with a formalization like:

1. $\Box_i(G), P$
2. $\Box_i(G \supset N), P$
3. $\neg G \supset \Box_i(\neg N), P$
4. $\neg G, P$
5. $\Box_i(N), P \ast 1, 1, 2$
6. $\Box_i(\neg N), MP, 3, 4$

Supposing that we are talking about moral obligations, this gives Jones a moral dilemma with two conflicting obligations. In our daily life, we have no problem coping with a situation like this without getting into a dilemma. But, Castañeda says, a principle like (K*) applies only when $A$ and $B$ are practitions. “If Jones is to visit Smith, Jones ought to notify Smith first” has in fact the form $\Box_i(g \supset N)$, where $g$ stands for the proposition “Jones is to visit Smith” and $N$ for the practition “Jones to notify Smith”, so the moral dilemma cannot be derived, because we never derive that Jones ought to notify Smith. In [10, p. 62f], Castañeda also discusses “the biconditional paradox”, which is very similar to the Chisholm paradox.

It is worth nothing, that $\Box_i(g \supset N)$ is equivalent to $g \supset \Box_i(N)$ in Castañeda’s system. Even in a minimal system without the act/circumstance distinction, such as SDL, we can use the formalization $g \supset \Box(n)$ to formalize the second premise without getting into a dilemma. Many, e.g. Fred Feldman, have
objected to such solutions, that \( g \supset \Box(n) \) is entailed by \( \neg g \), and it seems implausible that there is such a logical connection between the two premises [21, p. 97]. If this is a good objection, even Castañeda’s system is vulnerable to it. Given the basic assumptions of Castañeda’s system, it does not seem that the premise can be plausibly interpreted in a way not equivalent to \( \Box_i(g \supset N) \).

4.4.3 Lewis-style Systems

In Lewis-style systems, contrary-to-duty paradoxes are handled by means of the dyadic deontic operator.

Feldman formalizes the premises in the original Chisholm example as:

1. \( \text{MO}_{jt}(g) \)
2. \( \text{MO}_{jt}(n/g) \)
3. \( \text{MO}_{jt}(\neg n/\neg g) \)
4. \( \neg g \)

From the premises, it follows that, in Feldman’s system, \( \text{MO}_{jt}(n) \). This kind of inference is what Feldman calls “deontic detachment”.

There was a debate between Castañeda and Feldman, where Castañeda focused much on this point. He found it deeply counterintuitive that we can derive “Jones is not about to visit Smith and Jones ought to notify Smith”. In Feldman’s system, this means that Jones notifies Smith in the best worlds that are accessible to Jones, e.g., given the premises of the example, worlds where Jones really goes to help Smith, and this does not contradict the norm that in the actual world, where Jones does not go to help, the best thing would be to not notify Smith about coming. Feldman pointed out these things in a reply to Castañeda, and he also described a situation where it would be realistic to say that Jones in fact has a duty to notify Smith [22, p. 35]. Nevertheless, Castañeda stuck to his opinions in a further reply [12].

These special problems about inalterability arise because the action of notifying, whose obligatoriness is, in some sense, defeated, was to be performed before the action of visiting. In cases where our ideal obligation pattern has been altered because of our past wrongdoings, things are handled in a more straightforward way in a time-indexed deontic system such as MO, because we can then easily say that there is no question that the wrongdoing is something unalterable.

An example of this is Feldman’s treatment of the “second-best-plan paradox” [23, p. 325f]. The best thing a doctor can do might be to give a medicine \( A \) on Monday, and medicine \( B \) on Tuesday. But if the doctor gives medicine \( C \) on Monday, it would be bad to give \( B \) on Tuesday, and the best thing to do then would be to give medicine \( D \) instead. This doctor may have an absolute obligation of the type \( \text{MO}_{dt}(a \land b) \), from which \( \text{MO}_{dt}(a) \) and \( \text{MO}_{dt}(b) \) follow, but also a conditional obligation of the type \( \text{MO}_{dt}(d/c) \). At a later time, \( t_2 \), when \( c \) has already been given, the doctor instead has an absolute obligation like \( \text{MO}_{dt_2}(d) \).
Many examples of what is called “Contrary-to-Duty Imperatives” in the literature have the same structure as the Second Best Plan-example. For example, Lennart Åqvist and Jaap Hoepelman describe a “tree” system which they call DARB (deontic arbor, which is Latin for “tree”) in [4], and use this to handle a case with the man John, who operates a devilish machine with one button. If he never presses the button, nothing happens, but if he presses the button just one time, the whole earth blows up, which is the worst possibility, and if he presses the button two times, just America blows up, which is the second worst possibility. Initially, John ought to never press the button, but he actually presses the button at the next instant, so he ought to press the button again at the instant after that in order to save the earth, even if he cannot prevent the destruction of America. In DARB, we have the deontic operator SHALL, and the time operator, $\oplus$, which is interpreted as “it will be the case at the next instant that”. The deontic operator is defined in terms of the best possible world stories relative to a point of time. If $p$ stands for “John presses the button”, the premises can be represented as:

1. SHALL $\oplus \neg p$
2. $\oplus p$
3. $\oplus p \supset \oplus \text{SHALL} \oplus p$
4. SHALL($\oplus \neg p \supset \oplus \oplus \neg p$)

This is a solution to Second Best Plan-like examples, which is, in many aspects, similar to Feldman’s MO; one difference is that it does not use any dyadic deontic operator, even if there is one, Shall($p/q$), in the system. It may, like MO, work fine for such examples. However, if MO’s handling of cases like the version of the Chisholm example described above, when one ought to adopt one’s actions to one’s own future wrongdoings, is inadequate, I do not see how DARB could take us any further than MO, because the set of possible world stories narrows down with time, just as in Feldman’s system.

How can Castañeda’s and Feldman’s different intuitions regarding the Contrary-to-Duty examples be explained? There is often an ambiguity in the description of these examples. If it is assumed that the violation of a duty (or something that is a duty relative to some agent-time index $jt_1$), like going, is something unavoidable (relative to another index $jt$, where $t$ stands for an earlier time-point), because we cannot now do anything to prevent it, we should not say that the duty holds at $jt$, and we cannot derive any other duties at $jt$, such as the duty to notify, from that duty. The problem with MO is that it is assumed that the set of worlds accessible at a later time is always a (proper) subset of the accessible worlds at an earlier time, so it cannot, in my view, adequately represent situations when we have no control over what we will do at some later time.

This problem never occurs in CDM1, because the different indexed operators in that system are not interrelated on the basis of any notions about agents or points of time. One might argue for a strongly possibilist morality, where our duties now are determined by our optimal action pattern at all future times.
CDM1 is compatible with such a morality, but it does not force it upon is, and this seems to be right, given my ambitions, because such a morality can hardly be a conceptual truth. (This may not be a criticism against Feldman, because he explicitly states that he does not intend his MO principle as a conceptual analysis of moral obligation, but rather as a stipulative definition aimed to serve as a basis for a possibilist moral system [21, p. 207].) Another possible problem with claiming that our possibilities shrink with time is that it seems to exclude backwards causation by agents, where an agent at a time \( t \) may have greater possibilities to influence the events at an earlier time \( t_0 \) than the agent had at \( t_0 \). I do not think this should be excluded at an \( a \ priori \) level.

If the wrongdoing is not unalterable (relative to \( j t \)), we should, with Feldman, say that the original duties to visit and to notify before visiting still hold. What are we to do with Smith’s conditional obligation to not notify without visiting? With Feldman’s dyadic operator, we can formalize this as \( \text{MO}_{jt}(\neg n/\neg g) \). In CDM1, we have no dyadic operator. However, we can assume the conditional \( \bigcirc_{jt}(\neg g \supset \neg n) \), which in CDM1 does not trivially follow from \( \bigcirc_{jt}(g) \), and we cannot use it to do modus ponens in that situation. On the other hand, if the situation is such that not going is unalterable for Jones at \( t \), viz. if \( \Box_{jt}(\neg g) \) holds, \( \bigcirc_{jt}(\neg g \supset \neg n) \) may be be used to derive \( \bigcirc_{jt}(\neg n) \). In both respects, CDM1 offers the same resources as a system with dyadic operators, so the use of dyadic operators to handle Contrary-to-Duty situations seems to complicate more than it solves.

So, the intuitive plausibility of the different premises in the Chisholm example seems to be derived from different readings of the example. On Castañeda’s reading, the fact that Jones will not visit Smith is unalterable for Jones at \( t \), and on Feldman’s reading it is not. If we clearly distinguish between those readings, we have, as seen in an earlier chapter, no problem representing any of them in CDM1. If we mix them up, we will get into trouble no matter which deontic system we use.

4.4.4 Danielsson on Contrary-to-Duty Paradoxes

Sven Danielsson distinguishes between “circumstances” and “alternatives” [17, p. 13ff]. An alternative in Danielsson’s sense is a “still open possibility”, and to assert that \( a \) is an alternative, which we, in accordance with Danielsson’s convention in a later paper [19], may label \( \text{M}(a) \), is incompatible with asserting either \( a \) or \( \neg a \). He holds that \( \bigcirc(a) \), which he holds is about what we normally express with the Swedish present tense form of “ought to”, “bör”, which, according to Danielsson, is “prescriptive”, implies \( \text{M}(a) \). Any “paradox” about contrary-to-duty obligations can never arise, because holding that \( \bigcirc(a) \) is inconsistent with holding \( \neg a \).

I am not sure whether asserting “bör \( a \)” in Swedish always requires treating \( a \) as “open” in Danielsson’s strong sense. Danielsson acknowledges that we often need to say that someone has violated, or will violate an obligation, but
we should then, he says, use the imperfect form “borde a” instead of “bör a”. According to Danielsson, the contrary-to-duty paradoxes arise in modern English, because there is no such distinction with the English “ought to”. CDM1 is a classical logic with possible-world semantics, where it is assumed that every proposition has a determined truth-value at any world. The truth of \( \neg a \) certainly does not imply that \( a \) is impossible, in CDM1. Possibility is relative to indices, and \( \neg \diamond_i (a) \), for any index \( i \), entails \( \neg a \), but we may still have \( \diamond_j (a) \), relative to some other index, \( j \), as in one of the versions of the Chisholm Paradox discussed above.

Maybe, we use the distinction between “bör” and “borde” to announce a shift in alternative set in a speech context. The \( jt2 \)-duty to \( g \) might then be such that we in Swedish would express it with “Jones borde gå och besöka Smith [go to visit Smith]”. But even such duties imply that there is some index relative to which \( g \) is possible. Can we say that if \( i \) is the primary alternative set in a speech context, that which we hold implicit when we use the present tense form is such that when we assert \( a \), we also always ascribe to \( \Box_i (a) \), so we cannot claim “bör \( \neg a \)? My main problem with this proposal is that it seems to exclude e.g. certain possibilist moralities already at the conceptual level. Someone who ascribes to a strongly possibilist morality would not claim that there is any relevant “ought” such that Jones, in Holly Smith’s above-mentioned tunnel case, has no obligation to not change lanes nor accelerate, even though it is assumed that Jones is in fact going to change lanes. An actualist could, in such a case, announce the shift in alternative set by changing from “bör accelerera [accelerate]” to “borde varken accelerera eller byta fil [neither accelerate nor change lanes]” in Swedish. In English, the change would have to be announced in another way, perhaps by saying something like “Jones ought to accelerate, now that she is to change lanes, but she really ought to neither accelerate nor change lanes”.

4.5 Geach’s John/Tom-example

A world that is ideal relative to one context \( i \), need not be so relative to another context, \( j \), even though the world is possible relative to both contexts. In this way, we may handle examples like Peter Geach’s, where we want to say that Tom ought to be beaten up by John, but not that John ought to beat Tom up. The following statements seem to be analytically equivalent:

\((\text{BeatActive})\) John beats Tom up.

\((\text{BeatPassive})\) Tom is beaten up by John.

From \( \Box_i (\text{BeatActive}) \), we may then derive \( \Box_i (\text{BeatPassive}) \), and vice versa. Replacement of material equivalents still holds for \( \Box_i \), even though the consequence principle does not: any proposition logically equivalent with some-
thing that describes an open alternative also describes an open alternative. However, if the reason why we are inclined to claim that (BeatPassive) describes something obligatory, while (BeatActive) does not is that we think that John is not under any obligation to beat Tom, while Tom may be under an obligation to be beaten up by John, this can be handled by the distinction between different contexts. We may hold that $\Box_{Tom}(\text{BeatPassive})$ (and therefore $\Box_{Tom}(\text{BeatActive})$) holds, while $\Box_{John}(\text{BeatActive})$ does not. As long as we do not need to distinguish between saying that John is under an obligation that he beats Tom up, and that he is under an obligation that Tom is beaten up by him, we have no problem. I fail to see how this could be adequately handled in a system that does not contain indexed obligation/requirement operators.
5. Quantification in Deontic Logic

5.1 Introduction
Most discussion about deontic logic, as well as modal logic in general, has been in the context of propositional logic. However, if the basic ideas behind deontic logic are reasonable, it seems reasonable to require that we should be able to use the deontic operators also in a quantified language. Moreover, quantification may be directly necessary for the interpretation of the deontic categories. One example is statements about general rules. Such statements typically say that everybody under certain circumstances must make it the case that some entities have certain properties. It is hard too see how we could express such statements without a quantified language. In this chapter, I will first give a survey of some of the central issues in quantified modal logic. I will then go through some arguments regarding the extensionality of deontic contexts. If we could argue that deontic contexts are extensional in a sense in which ordinary modal contexts are not, we would have a case for not treating quantified deontic logic as a branch of quantified modal logic. I will argue that the arguments for the extensionality of deontic contexts fail, and that the problems related to quantification in deontic contexts should thus not be treated as specific to deontic logic any more than, say, problems of interpreting belief statements occurring inside the scope of deontic operators.

5.2 General Issues in Quantified Modal Logic
There are some problems in modal predicate logic in general concerning the interpretation of the individual domain. Shall we have a fixed domain with the same individuals for all worlds, or shall we use varying domains, where an individual that exists in one world need not exist in another? Are we to have rigid designators, where an individual term refers to the same individual in every world, or non-rigid designators, where this is not necessarily the case?

It is not my aim to go further into these problems here, even though they sometimes have been discussed specifically in the context of quantified deontic logic. For a discussion about systems of quantified deontic logic with both fixed and varying domains, and rigid and non-rigid designators, I refer to Gerhard Schurz’s treatise *The Is-Ought Problem* [53] on Hume’s Law and deontic logic. Stig Kanger, in “New Foundations for Ethical Theory” [38], proposed a system of quantified deontic logic with fixed domain and non-rigid desig-
nators. J.A. van Eck has, like Schurz, in his “A system of temporally relative modal and deontic predicate logic” [58], proposed combined modal and deontic predicate logics: the system of quantificational modal temporal logic (QMTL), and its extension QDTL, which also contains a dyadic deontic operator. In addition, his systems are temporally relative, so that the truth of a formula is time-relative. He has rigid designators and a world-relative domain.

5.3 Extensionality as a Case for Special Treatment of Quantified Deontic Logic

5.3.1 Castañeda’s Extensionality Argument

Hector-Neri Castañeda has argued that deontic judgments are extensional, in the following senses:

1. If it is true e.g. that Richard Nixon ought to finish the Vietnam War in 1971, then it is also true that the man who in 1925 was $P$ ought to finish the Vietnam War in 1971, if “the man who in 1925 was $P$” is any definite description denoting Nixon.

2. If an agent has a duty to do $A$ to some individual $c$, then that agent also has a duty to do $A$ to any $d$, such that $c = d$ [9, p. 231].

Because ordinary propositional modalities are intensional, Castañeda’s argument is relevant to the question whether quantified deontic logic should be treated as a branch of quantified modal logic.

Castañeda’s claims become problematic if we consider deontic statements with identity statements inside the scope of the deontic operator. Gerhard Schurz illustrates the point that individual terms cannot be rigid if they are understood as definite descriptions with an example, where the boyfriend of Susan actually is identical with the husband of Mary [53, p. 192f]. We might well say that it ought to be that these are not identical. But other identity statements involving “Mary’s husband”, or “Susan’s boyfriend”, might describe some desirable or neutral state. In such a case, a pair of identity statements like $m = s$ and $m = t$ might be true, where $m = s$ stands for something desirable, and $m = t$ stands for something undesirable. We might then hold that $\neg \Box (m = s)$, but also $\Box (m = t)$ are true, even though $s = t$ is also true.

Nevertheless, it seems that Castañeda’s observation is right in the majority of cases one can think of. In some cases, there is an ambiguity in the interpretation of a definite description. Assume that Mary’s husband ought to end his affair with Susan immediately, and also ought to apologize to Mary later for his infidelity. Is it true that Susan’s boyfriend ought to apologize to Mary tomorrow? Yes, if we interpret the description as “the person who is now Susan’s boyfriend”, but it sounds strange if we interpret it as “the person who is at that later time Susan’s boyfriend”. The same ambiguity may, of course, be found in descriptive statements, like “Susan’s boyfriend will apologize to
Mary”. But one difference may be that we, in the deontic case, do not consider the identity of Susan’s future boyfriend in the actual world, because it is still open to Mary’s husband whether he will continue his affair with Susan.

So the problem seems to be related to the fact that some states are regarded as open alternatives. We cannot suppose that a person who ought not to continue to be Susan’s boyfriend also, in fact, will not continue being that, and we cannot then always substitute definite descriptions in open alternatives. But the fact that Nixon is the one and only man who in 1925 was P should presumably be seen as something fixed when we discuss his obligations in 1971. Therefore, we may safely substitute “the man who in 1925 was P” for “Richard Nixon” in Castañeda’s example.

5.3.2 Substitution of Identical Individuals and the Jephta Dilemma

Other alleged examples of paradoxical results arising from allowing substitution of identical individuals in deontic contexts, which could be seen as giving extra arguments against the extensionality of these contexts, have been discussed by some writers. The so-called Jephta Dilemma, inspired by Book of Judges, was introduced in the debate by von Wright, in [64, p.118f], and discussed in more detail in [63, p.79]. Von Wright, writing in the context of a dyadic propositional deontic logic, describes the story as a “predicament”, or dilemma: Jephta has promised the Lord to sacrifice the first being he will meet on his return home, and we may assume that he ought to do whatever he has promised the Lord. But the first being he meets on return is his own daughter, Miriam, and he ought not to kill his own daughter. Von Wright claims that Jephta’s dilemma arises because he has given a promise he should never have given, and he posits a theorem, stating that an inconsistent obligation can only arise when something forbidden is the case, which is called the “Jephta theorem”.

\[
\text{JephtaTheorem} \quad \circ (\neg \top / p) \supset \circ (\neg p)
\]

Lennart Åqvist relates the Jephta Dilemma to the question about the validity of substitution of identical individuals in deontic contexts. He states the premises like this.
1. Miriam, Jephta’s daughter, is the first being that will meet Jephta on his return home.
2. Jephta ought to immolate the first being that he will meet on his return home.
3. So, Jephta ought to immolate Miriam.

Åqvist, who apparently wants to avoid a dilemma in a situation like this, says that it seems that “the argument must be invalid” [3, p. 82]. He also says that a system like Kanger’s would block “deontic analogues of the Morning
5.3.3 Forrester and the Multiplied Beneficence Argument

James W. Forrester also discusses problems related to the extensionality of deontic contexts. He discusses the following principles [26, p. 301]:

\[ QA1 \ \forall x F(x) \supset F(x) \]

\[ QA2 \ \exists x F(x) \supset F(x) \]

QA2 cannot hold in deontic contexts, if we interpret the existence quantifier as “actually exists”. He has an example with Christopher, who ought to build St. Swithin’s Cathedral [26, p. 301f]. This, Forrester says, does not imply that the cathedral will exist at any time in the actual world. I do not share his intuition here. We might, of course, say that Christopher ought to build a cathedral at a place, even if we do not know whether he will do so, but we would not say, I think, that somebody ought to do something with a specific, named, cathedral if we do not hold that this cathedral actually exists. Therefore, I think, it is an anomaly to say that somebody ought to build a specific cathedral. I grant that we may evaluate the existence of certain individuals morally, but then we usually refer to existing individuals, and say things like “This person should
never have been brought into existence”, or “Christopher should never have built St. Swithin’s Cathedral”.

Forrester uses the alleged non-extensionality of deontic contexts to argue against what he calls “the multiplied beneficence argument” (MBA) [26, p. 188–96]. In his example of an MBA-structure argument, it is assumed that Norma has a duty to help a specific child who is starving. But there are millions of other children in similar circumstances. Norma can help any of these, and it seems that her duty to help the first could be multiplied into a duty to help all the children. Forrester points out that one might fulfill general duties by doing things that are not duties in themselves. If we assume that Norma ought to help \( n \) starving children, she may discharge this duty by helping some group of \( n \) children. This does not mean that she has any particular duty to that specific group. The intensionality of deontic contexts, in Forrester’s sense, means that even if we allow that Norma ought to help all the starving children, we cannot infer that there are specific children Norma ought to help [26, p. 192], i.e. QA1 does not hold.

According to Forrester, this refutation of MBA by means of deontic logic may be of great philosophical and practical importance. He implies that Peter Singer’s argument to the effect that we have far-reaching duties to alleviate suffering in foreign parts of the world is an example of MBA. He also claims that MBA reasoning has, for example, made legislators unwilling to recognize minimal duties of beneficence, because they fear it will lead to slippery slopes, where it becomes impossible to limit the duties in any reasonable way.

Is Peter Singer’s argument really an example of MBA? Forrester refers to his well-known article, “Famine, Affluence, and Morality”. In this article, Singer assumes that “if it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it” [55, p. 231]. He then argues that it is possible for affluent people to alleviate world famine by giving away a lot of their resources, without thereby sacrificing anything that reasonably can be said to be of equal moral importance. It is hard to see why this should be an MBA. Singer’s example, where he says that he would have a duty to help a drowning child out of a shallow pond, even if he then would get his clothes muddy, may give some impression of an MBA. Suppose someone argued that you ought to help this child because by doing so you can save its life at minimal cost to yourself; there are many other children you can help at minimal cost; therefore, you ought to help all of these children. This would be an MBA, and, as Forrester says, would give you an obligation that is not only heavily demanding, but actually impossible to fulfill. But Singer only gives the drowning child-case as an example of an application of his general principle about the prevention of evil. (In the drowning child-case, we might say that we really ought to help this particular child, because in doing so we do not diminish the resources we have left to help others, unlike when we give away money to help one particular child on a list, as in Forrester’s example.)
Forrester does not give any other really convincing examples of MBA reasoning in philosophical or other contexts. He quotes some British law commissioners in India who say that they are unable to see where they can draw the line, if they make someone legally punishable for letting “a fellow creature die of hunger at his feet”, while having abundance of wealth. This may, as Forrester implies, be based on MBA reasoning, but I find this unlikely, because the commissioners seem to recognize moral duties of beneficence, because they call the uncharitable person “a bad man”, and the MBA would, if valid, lead to a slippery slope in the case of moral, as much as legal, duties. So, if the commissioners believed in MBA reasoning, they would probably also be unwilling to recognize such a moral duty.

In short, Forrester’s argument for the utility of deontic logic should be viewed with suspicion. His refutation of MBA also seems to be of limited interest to deontic logic. In order to defeat the MBA we do not have to assume that QA1 is invalid; it is only required that we do not have what may be termed the existential analogy of the converse Barcan formula as a theorem:

\[ \text{ExcBF} \quad \square \exists x A \supset \exists x \circ A \]

No system of deontic predicate logic I can think of has that.

The rejection of QA1 also seems to give implausible results. If we assume that Norma really ought to help all children on the list, it seems that we are, under normal circumstances, entitled to draw the conclusion that if Tomas is on the list, she ought to help Tomas.

Again, this seems to point to the need of distinguishing between alternatives and fixed circumstances. We should be entitled to draw the conclusion, if the existence of Tomas is not an alternative for Norma.
6. Supererogation and Dilemmas

6.1 Introduction
Some philosophers claim that we sometimes may face moral dilemmas, traditionally characterized as situations where we cannot avoid breaking at least one moral norm, or that there exists supererogation, or other deontic categories going beyond the standard classification of acts or states of affairs as obligatory, forbidden or indifferent. If so, these are two aspects of moral thought that SDL simply does not allow for, for reasons explained in the introduction (sections 1.5.5–1.5.6), and CDM1 is not designed to handle them either. As explained in chapter 2, the (DPos) schema yields a contradiction in case of a situation where we have two norms that are not jointly satisfiable relative to a context. Even some of those who believe that there are no real instances of dilemmas or supererogation, perhaps because they believe that e.g. some kind of consequentialism is the true moral theory, may think that there is nothing logically inconsistent with a moral system that allows for these things. In this chapter, I will go through some proposals for handling these alleged features of moral reasoning.

I do not intend to give any comprehensive picture of this area, but I will give examples of how these problems might be handled in the framework of propositional logics relatively close to ordinary modal logic, and contrast these approaches with some alternative proposals, like Chisholm’s agency-based treatment of supererogation and Routley’s and Plumwood’s relevant treatment of dilemmas. In the end, I will also discuss how CDM1 might be modified in order to handle them, if that is found necessary. In doing this, I do not want to defend the existence or possibility of neither dilemmas nor supererogation; my aim is rather to investigate into how much an approach to deontic logic of the type sketched in chapter 2 has to be modified in order to take care of the intuitions among people who believe in those features.

6.2 Chisholm on Supererogation
6.2.1 Chisholm’s Nine-Category Proposal
In his article “Supererogation and Offence”, Roderick Chisholm presented a schema according to which actions are classified according to the goodness, badness or neutrality of their performance and non-performance. This
gives $3^2 = 9$ categories of actions to which he gives the following names, where $G(B,N)((\sim)A)$ represents the goodness (badness, neutrality) of the (non)-performance of $A$ [15, p. 427].

1. $A$ is totally offensive: $B(A) \land B(\sim A)$

2. $A$ is an offence of commission: $B(A) \land N(\sim A)$

3. $A$ is forbidden: $B(A) \land G(\sim A)$

4. $A$ is an offence of omission: $N(A) \land B(\sim A)$

5. $A$ is totally indifferent: $N(A) \land N(\sim A)$

6. $A$ is a supererogatory omission: $N(A) \land G(\sim A)$

7. $A$ is obligatory: $G(A) \land B(\sim A)$

8. $A$ is a supererogatory commission: $G(A) \land N(\sim A)$

9. $A$ is totally supererogatory: $G(A) \land G(\sim A)$

According to Chisholm, it is possible to define these different concepts in terms of “ought-to-be” [15, p. 428f]. An act is good to perform if the situation “which the act would bring about” ought to exist, it is bad to perform if that situation ought not to exist, and it is neutral if it is neither good nor bad. Chisholm does not work out any formal system for dealing with his nine categories, but he has some propositional operator in mind. Because Chisholm bases his definitions on evaluative notions, it would appear natural to ground a formal treatment of these notions in some kind of preference logic. However, it might still be of interest to see how many of his categories that can be retained within the framework of traditional deontic logic. It is clear that if we want to keep the nine categories, we cannot just add a traditional $\Box$-operator to propositional logic, interpret the situation corresponding to the non-performance of an act $A$ as the negation of a proposition $a$ stating that the act is performed, $\neg a$, and interpret that $a$ ought to be as $\Box(a)$, and that $a$ ought not to be as $\Box(\neg a)$. $\Box(\neg a)$ is in classical logic logically equivalent to $\Box(\neg a)$. This would mean that if the non-performance of the act is bad, its performance would be good, and vice versa, which would exclude all categories where just one member in a performance/non-performance-pair is neutral, as well as the two categories where both members have the same non-neutral status. What would remain would be the three traditional categories of the obligatory, the neutral, and the forbidden. Chisholm himself says that “non-performance should not be said to entail performance of non-performance; deliberately refraining from showing gratitude would exemplify performance of non-performance” [15, p. 423], which might be offensive when the mere non-performance is not.

Can the addition of an agent-relative “see to it”-operator in the style of e.g. John Horthy [35, p. 14] do the job Chisholm wants? That the performance of $A$ by $x$, written $xcstit : a$, where $a$ is some proposition that is realized by the action, is good would be interpreted as $\Box(xcstit : a)$. That the non-performance is good would be $\Box(\neg(xcstit : a))$. That the performance is bad would be $\Box(xcstit : \neg a)$; and this would be equivalent to saying that the performance of the non-performance of $A$ by $x$ is good. But if we assume that the consequence principle holds, the badness of the performance would here entail
the goodness of the non-performance, because \( xcst : \neg a \) entails \( \neg (xcst : a) \), which would exclude the totally offensive and the offences of commission. Even if we reject the consequence principle, as I do in the system presented in chapter 2, it still holds that with a principle against dilemmas of the form \( \bigcirc (xcst : a) \land \bigcirc (xcst : \neg a) \), the totally supererogatory would be excluded. Not everyone that welcomes supererogation would also welcome dilemmas, I suppose.

Perhaps, it would be a better idea to use some kind of permissibility operator when we try to interpret the goodness and badness of states of affairs. I have not found out any treatment that can capture all Chisholm’s categories. For example, if we interpret the badness of \( a \) as \( \neg Pa \), we cannot have that both \( a \) and \( \neg a \) are bad, because \( \neg Pa \) implies \( P\neg a \) in ordinary deontic logic.

If we try to use some kind of preference logic in order to represent the categories, the categories of the totally offensive and totally supererogatory both directly contradict what has been called a condition of “non-duplicity”, which is commonly assumed to hold for monadic value predicates like “good” and “bad”. According to Sven Ove Hansson’s definition, a monadic predicate \( H \) satisfies this condition iff for all \( p: \neg (Hp \land H\neg p) \) [32, p. 117]. In a later paper, Chisholm himself, and Ernest Sosa, proposed defining the deontic concepts in terms of a logic for intrinsic value that satisfies the non-duplicity condition for goodness and badness. Chisholm and Sosa do not comment on the exclusion of the categories 1 or 9 in [15], nor do they comment on Chisholm’s different approach to supererogation presented in [14], which I discuss below in section 6.2.2. McNamara’s DWE-logic is a kind of deontic logic based on a preference relation I will discuss in section 6.3.

How plausible are Chisholm’s nine categories? As he says, we can distinguish between \( 2^9 = 512 \) different types of moral systems in regards to which categories they exemplify. Chisholm says that “the strict utilitarian may be able to find instances of each of the odd-numbered possibilities” [15, p. 427]. This is interesting, because utilitarianism, and consequentialism in general, is often thought of as a theory that classifies every action into the three traditional categories of obligatory, indifferent, and wrong. Chisholm gives an example of a “totally supererogatory” act in a consequentialist system. If the only morally relevant consequences were that its performance would give one person a certain amount of pleasure, and its non-performance would give another person the same amount of pleasure, it would be “totally supererogatory” in Chisholm’s sense. Most consequentialists would classify such an act as permitted, but not obligatory. They would of course agree that the world brought about by the performance, or non-performance, of such an act would be good in an evaluative sense, but that would not, according to the consequentialists, be a moral evaluation of the act.

As for the totally offensive, Chisholm says that this is also exemplified in utilitarianism, if we replace “pleasure” with “displeasure” in the example of the totally supererogatory. The consequentialist’s answer would be the same:
both acts would be permitted but not obligatory. Chisholm’s category of the
totally offensive seems like an example of a moral dilemma where we cannot
avoid doing the bad thing. Another of his examples of the totally offensive
is more like a classical dilemma: we assume that promise-breaking is always
bad, and that we have promised a contradiction, such as to go and not to go.
The act of going is then something totally offensive.

6.2.2 Chisholm’s Agent-Relative Proposal

In a somewhat later article, “The Ethics of Requirement”, Chisholm claims
that “all the fundamental concepts of ethics” can be defined using “p requires
q”, written \( pRq \), as the single primitive, where \( p \) and \( q \) are propositional vari-
ables [14, p. 147]. These requirements may be overridden. He then defines
that it ought to be that \( q \) (written \( \Box(q) \)) as saying that there is a requirement
for \( q \) which has not been overridden. He also claims that “\( S \) ought to do \( q \)”
can be defined by means of agent-relativization (in the style of Kanger above)
as “It ought to be that \( S \) bring it about that \( q \)”, written as \( \Box(Sq) \). He proposes
a definition of supererogation by omitting the reference to the agents bringing
about in the scope of the obligation operator, and he uses the agent-relative
operator also in his definition of other deontic categories.

1. \( S \)’s doing \( q \) is supererogatory: \( \Box(q) \land \neg(Sq) \land P(Sq) \)
2. \( S \)’s doing \( q \) is offensive: \( \Box(\neg q) \land \neg(Sq) \land P(Sq) \)
3. \( S \)’s doing \( q \) is indifferent: \( P(q) \land \neg(q) \land P(\neg(Sq)) \land P(Sq) \)

Can this idea handle all categories Chisholm himself proposed in [15]? If
he wants to include the totally offensive or totally supererogatory, this seems
to require some “non-classical” interpretation of the ought-to-be-operator, in
order to permit that both \( q \) and its negation are obligatory in the ought-to-
be-sense. In his later article, he says that the “concepts of supererogation and
offence have a place, of course, only in an ethic which is latitudinarian”, and
that they have “no place, presumably, in strict utilitarianism” [14, p. 153]. He
notes the difference between this analysis of the concept of offence and the
one in [15], and says that his new analysis seems “clearly preferable” [14, p.
153]. He does not discuss his earlier examples of plausible applications of, for
example, the concept of the totally offensive. As I said above, his utilitarian
examples do not seem to have much to do with the utilitarian conception of
moral worth, but his promise-giving examples may be more plausible with
respect to certain moral systems; they seem to capture much of what is often
meant by “moral dilemmas”. But the logical means for handling dilemmas
will be discussed later in this chapter.
6.3 McNamara’s View on Supererogation

Paul McNamara claims to have a “logic for common sense morality” in [43]. In addition to the obligation operator $\Box(a)$, he adds three other unary operators, $\text{MI}(a)$, read as “doing the minimum” morality requires involves seeing to it that $a$ [43, p. 178], $\text{MA}(a)$, read as “doing the maximum” morality allows, what is morally best among the morally acceptable options, involves seeing to it that $a$ [43, p. 179], and $\text{IN}(a)$, read as that it is morally indifferent whether $a$ or $\neg a$ [43, p. 171]. Permissibility and impermissibility are defined as usual. He also has the following definitions:

1. It is gratuitous that $a$: $\text{GR}(a) \iff \neg \Box(a)$
2. It is optional that $a$: $\text{OP}(a) \iff \neg \Box(a) \land \neg \Box(\neg a)$
3. It is morally significant that $a$: $\text{SI}(a) \iff \neg \text{IN}(a)$
4. It is supererogatory that $a$: $\text{SU}(a) \iff P(a) \land \text{MI}(\neg a)$
5. It is suboptimal that $a$: $\text{SO}(a) \iff P(a) \land \text{MA}(\neg a)$

McNamara provides a “Doing Well Enough” (DWE)-semantics, which is like the standard Kripke semantics, with an accessibility relation, $A$, and which also contains a point-relative ranking, $\leq$ of points, which is transitive and holds between the worlds that are accessible relative to a point (viz. for any points $i, j, k, k \leq i \land j$ or $j \leq i \land k$ iff $A(i, j) \land A(i, k)$). The truth-conditions for obligation are defined as usual. $\text{MA}(a)$ is interpreted so that there is an accessible point such that $A$ is true in every point that is ranked at least as high. $\text{MI}(a)$ is interpreted so that there is an accessible point such that $a$ is true for every point that is ranked at least as low. $\text{IN}(a)$ is interpreted so that there is, for every accessible point, one equally ranked point where $a$ is true, and one equally ranked where $\neg a$ is true.

McNamara also has a minimal normal DWE-logic that has been proven sound and complete with respect to the class of DWE-models. Except for the addition of the extra operators, this is much in the spirit of standard deontic logics, with a necessitation rule for all the operators $\Box$, $\text{MA}$, and $\text{MI}$ ($\Box/\text{MA}/\text{MI}(\top)$ holds), and two axioms, $\Box(a) \supset (\text{MI}(a) \land \text{MA}(a))$, and $(\text{MI}(a) \lor \text{MA}(a)) \supset P(a)$, from which the standard deontic axiom that everything obligatory is permitted obviously follows. So his theory is, as it stands, vulnerable to Samaritan Paradoxes and other SDL problems discussed in 1.5.

We may note that the DWE-semantics is, in its relative ranking of worlds, similar to the semantics used in dyadic systems, such as that of David Lewis. If we would like to incorporate a DWE-type approach to supererogation into a Lewis-style system, we could use two different rankings, one “rough ranking” ($\succ_R$), which determines the borders for what is acceptable, given some proposition, and one “fine ranking” ($\succ_F$), which gives room for the more fine-tuned deontic categories. They would be related so that if $i \succ_R j$, then $i \succ_F j$.

“$a$ ought to be the case given $b$” ($\Box(a/b)$) could be defined as “$a$ is true in all $\succ_R$-best worlds where $b$ is true”. $\text{MA}(a/b)$ could be defined as “$a$ is true
in all \( >_F \)-best worlds where \( b \) is true”. \( \text{MI}(a/b) \) could be defined as “\( a \) is true in all the \( >_F \)-worst of the \( >_R \)-best worlds where \( b \) is true”. \( \text{IN}(a/b) \) could be interpreted so that for every one of the \( >_R \)-best worlds where \( b \) is true, there is one equally \( >_F \)-ranked world where \( a \land b \) is true, and one equally ranked where \( \neg a \land b \) is true.

In any case, McNamara’s handling of the supererogation problem is more complicated than Chisholm’s in [14] in that it requires more primitive deontic operators. On the other hand, it is simpler in that it does not require a deontic language where ought-sentences both with and without a see-to-it-operator are allowed for.

We might compare McNamara’s framework with Chisholm’s proposals. Can Chisholm’s nine categories be captured in the DWE framework? I do not see how the first and the last category, the totally supererogatory and the totally offensive, could be captured. As we have seen, these categories could not be captured in standard preference logic of the kind Chisholm himself later endorsed in [16]. However, we can find a natural correspondence with the seven remaining categories if we let \( a \) denote the proposition stating that \( A \) is performed.

1. \( A \) is an offence of commission: \( \text{SO}(a) \)
2. \( A \) is forbidden: \( \bigcirc(\neg a) \)
3. \( A \) is an offence of omission: \( \text{SO}(\neg a) \)
4. \( A \) is totally indifferent: \( \text{IN}(a) \)
5. \( A \) is a supererogatory omission: \( \text{SU}(\neg a) \)
6. \( A \) is obligatory: \( \bigcirc(a) \)
7. \( A \) is a supererogatory commission: \( \text{SU}(a) \)

One might argue that the DWE-logics yields problematic results in the following type of case\(^1\). Let \( m \) stand for the proposition that I give Johan money, and let \( u \) stand for the proposition that I am unkind to Johan. Suppose that there are four accessible worlds, where each combination of \( m \) and \( u \) is realized: \( w_1 \models m \land \neg u, w_2 \models m \land u, w_3 \models \neg m \land \neg u, w_4 \models \neg m \land u \). Assume that \( w_1 \) is the best world, but that \( w_2 \) and \( w_3 \) are also acceptable and equally ranked, and that \( w_4 \) is unacceptable. In the best state, I give money and also avoid being unkind; however, I do not have to do either. But if I am unkind or do not give money, I have to compensate by avoiding the other. In DWE, \( \text{SU}m \) and \( \text{SU}\neg u \) would come out as false. E.g. \( \text{MI}\neg m \), which is implied by \( \text{SU}m \), is false, because doing the minimum morality requires may involve seeing to it that \( m \), if I am unkind. But one might still hold that \( m \) is something supererogatory, presumably because it is something morally good that is not unconditionally morally required. One possible answer would be to say that \( m \) is not unconditionally supererogatory any more than it is unconditionally required; what is supererogatory is the combination \( m \land \neg u \). States like \( m \) may be captured by

\(^1\)This was pointed out by Johan Gustafsson.
some weaker concept of morally significant non-minimality defined in terms of the operators in DWE:

\[(SNM) \quad SNM(a) \leftrightarrow_{\text{def}} OP(a) \land SI(a) \land \neg MI(a)\]

In the situation described above, SNM(m) holds, because m is not in itself required and not necessarily involved in fulfilling the minimal requirements of morality, and we do not, for every world, have an m-world and an \(\neg m\)-world that is equally ranked. SU(a) implies SNM(a) in DWE, but the converse does not hold (see McNamara’s partition in [43, p. 193]).

Chisholm did not provide a fully-worked out system for his different conceptions of supererogation, and it might not be possible to give a definite answer to what his response to cases such as this would have been (at a given moment). With his value-based approach in both [15] and [16], treating e.g. the money-giving as supererogatory, would require that we regard the money-giving as good, which seems reasonable, and the refusal to give money as neutral, which seems doubtful; it does not seem reasonable to say that the refusal is neutral if I am unkind. With the agency-based approach in [14], we would have to regard the negation of my seeing-to-it that I give money as unconditionally permitted, which seems implausible.

6.4 Deontic Dilemmas and Classic Logic

6.4.1 Schotch’s and Jennings’s Non-Kripkean Deontic Logic

Peter K. Schotch and Raymond E. Jennings present, in their article “Non-Kripkean Deontic Logic” [52], a system, which they claim can be used to cope with the possibility of moral dilemmas. They want to allow dilemmas of the form \(\Box(a) \land \Box(\neg a)\), but not of the form \(\Box(a \land \neg a)\), and, on the syntactical level, they do this by rejecting the derivability of the latter from the former, by rejecting the (K) schema. In this way, we avoid the so-called deontic explosion, viz. we cannot derive \(\Box(b)\), for any b, from \(\Box(a) \land \Box(\neg a)\), because this requires deriving \(\Box(a \land \neg a)\), and then, with the consequence principle, \(\Box(b)\). They present two alternative semantical structures for this non-normal logical system. In the first, they allow multiple sets of accessible worlds for one possible world, and \(\Box a\) is said to be true iff a is true in all worlds in at least one such set. In the second, they have one accessibility relation, but hold that \(\Box a\) is true in a world iff a is true in at least one accessible world.

Is it not clear that \(\Box(a)\) in their second semantics is equivalent to \(\Diamond(a)\) in the minimal normal modal logic, K? \(\Diamond(a)\) is also defined as truth in at least one accessible world, and K also contains no specific restrictions on the accessibility relation present in the stronger modal systems. If they define the concept of permissibility, P(a) as \(\neg \Box(\neg a)\), this becomes equivalent to \(\Box(a)\) in K. So, if we, contrary to what is normally assumed, state that \(\Diamond\) is to be
interpreted as \( \bigcirc \), and \( \Box \) is to interpreted as \( P \), \( K \) can do the job of Scotch’s and Jennings’ system. This would be a simple move.

Unfortunately, there are serious problems with this approach. In his article, “A logic for deontic dilemmas”, Lou Goble develops a new approach to the problem moral dilemmas pose for deontic logic, while criticizing solutions like Scotch’s and Jennings’ [29, p. 466f]. The main goal he has in mind is to avoid deontic explosion while keeping the validity of intuitively valid inferences like the argument below, which we denote (Service):

1. I ought to fight in the army or perform an alternative service for my country.
2. I ought not to fight in the army.
3. So, I ought to perform an alternative service.

A problem with many alleged solutions to the problem of deontic dilemmas is that they fail to validate the argument (Service). In SDL, (Service) could be formalized like:

1. \( \bigcirc(f \lor s) \)
2. \( \bigcirc \neg f \)
3. \( \bigcirc[(f \lor s) \land \neg f] \lor \land \bigcirc d \)
4. \( \therefore \bigcirc s \lor \bigcirc \land \bigcirc d \)

We use the principle \( (\land d) \) with replacement of material equivalents. In some deontic systems that are designed to handle dilemmas, like Scotch’s and Jennings’, this principle is, as we have seen, rejected. It seems that the first to point out this type of problem was Bas van Fraassen in [59, p. 18].

6.4.2 Goble’s Proposal

Goble’s proposal is to change traditional deontic logic into what he calls a DPM-logic (Deontic Logic with Permitted Inheritance) by restricting the consequence principle to a Rule of Permitted Inheritance

\[ \text{(RPM)} \quad \text{If} \quad \vdash a \supset b, \text{then} \quad \vdash \bigcirc a \supset (\bigcirc a \supset \bigcirc b) \quad [29, \text{p. 473}] \]

In his DPM-systems, he also excludes the principle (D) in standard deontic logic, which says that everything obligatory is permitted. In the system DPM.1, he wants to allow for dilemmas of the form \( \bigcirc(a \land \neg a) \), and so excludes the axiom

\[ \text{(P)} \quad \vdash \neg \bigcirc \bot \]

He also offers a second version, DPM.2, which includes (P), but still allows for dilemmas of the form \( \bigcirc a \land \bigcirc(\neg a) \). In this system, he has to replace the principle of deontic conjunctive introduction, which he calls (AND), viz. that \( \bigcirc(a \land b) \) follows from \( \bigcirc(a) \land \bigcirc(b) \), in standard deontic logics with the weaker

\[ \text{(PAND)} \quad \vdash \bigcirc(a \land b) \supset [(\bigcirc a \land \bigcirc b) \supset \bigcirc(a \land b)] \]
At last, he offers the version DPM.3 with (P AND), but without (P). It is easy to verify that arguments like (Service) are indeed validated in the DPM-systems.

Goble has also proved that his DPM-systems are sound and complete with respect to different variants of neighborhood semantics, similar to the semantics I propose for the $\Box$-operator in CDM1, in 2.

### 6.5 A Relevant Approach: Routley’s and Plumwood’s Deontic Logic

In the paper “Moral dilemmas and the logic of deontic notions”, Richard Routley and Val Plumwood specify a deontic logic by adding the $\Box$-operator to relevant logic. They retain the consequence principle, but of course with relevant entailment instead of material implication. As they point out, including (D) leads to inconsistency but not triviality, because of the invalidity of ex falso quodlibet, if moral dilemmas are accepted, so they want to exclude (D). They claim that their system avoids the deontic paradoxes that result from the consequence principle, which they claim come from “reliance of material and strict conditionals, as an adequate representation of implication”, and may be avoided by using relevant entailment instead. They claim that the Good Samaritan Paradox can be avoided, because the material implication in a conditional statement like “We have an obligation to feed the starving poor $\supset$. There exist starving poor. […] cannot be upgraded to an entailment” [49, p. 23]. This seems like a strange concept of entailment: there is one interpretation of (S) “We have an obligation to feed the starving poor” which is “We have an obligation to feed any existing starving poor” and does not imply the existence of starving poor, but this could be handled without any special concept of entailment, and there is also another natural interpretation of (S) that clearly contradicts the statement that there are no starving poor. When confronted with other Good Samaritan-like cases, they say that there exists an entailment, but that it should not be considered a logically provable entailment [49, p. 24]. Their example is the “Paradox of the Repenter”, where “Grannie repented for having killed the vicar” implies “Grannie killed the vicar”.

We may note that Routley and Plumwood want to retain the derivation behind the alleged Alf Ross Paradox. Some non-classic logics might want to argue against this on the ground that they do not accept the derivation from $a$ to $a \lor b$ as universally valid; they want it to be valid only if the concept of $b$ is contained in the concept of $a$. Routley has criticized these attempts in [50, p. 97–100]. Moreover, it is hard to see how they could be relevant in avoiding the Ross Paradox. The concept of $b$ is certainly contained in the concept of $a \land \neg b$, and the derivation from $\Box(a \land \neg b)$ to $\Box[(a \land \neg b) \lor b]$ is certainly an instance of the Ross Paradox, if anything is.
Let us take a look at the semantics behind Routley’s and Plumwood’s relevant system, to see how dilemmas and the entailment relation are handled. For their basic relevant logic, they use semantic models of the form \(\langle T, K, 0, R, *, v \rangle\), where \(\langle T, K, 0, R, * \rangle\) is a model structure, \(T\) is the actual world, \(K\) is a class of worlds or situations, \(0\) is the class of so-called regular worlds, where all theorems hold, \(R\) is a three-place relation on elements of \(K\), \(\ast\) is a one-place “reversal” operation on \(K\) used to interpret negation so to avoid disjunctive syllogism and ex falso quodlibet, and \(v\) is a valuation function, which assigns the values 0 or 1 to sentential parameters, or atomic formulae. \(R(a, b, c)\) can be read as saying that “\(b\) and \(c\) are pairwise accessible from \(a\), or . . . \(a\) and \(b\) are compatible relative to \(c\)” (I think they intend “\(c\) and \(b\) are compatible relative to \(a\)” [50, p. 299f]. If there is at least one regular world \(x\), such that \(R(x, a, b)\), every \(p\) valued to 1 in \(a\) is valued to 1 in \(b\) too: \(a\) is included in \(b\), which they write \(a \leq b\).

Truth is related to \(T\), the actual world, so that a formula \(a\) is said to be true in a model iff \(a\) is interpreted as having the value 1 in \(T\) (written \(I(a, T) = 1\)), \(a\) is valid in a model structure iff it is true on all its valuations, and \(a\) is logically valid in a system iff it is valid in all model structures for the system [50, p. 302]. The R-relation is central to their interpretation of entailment. An entailment between \(a\) and \(b\) is interpreted as having the value 1 in a world, \(w\) (written \(I(a \rightarrow b, w) = 1\)), iff for every \(x, y \in K\), if both \(R(w, x, y)\) and \(I(a, x) = 1\) then \(I(b, y) = 1\). Conjunction and disjunction are interpreted as usual. Negation is interpreted so that \(I(\neg a, w) = 1\) iff \(I(a, w^\ast) \neq 1\).

This is the semantics for the most basic relevant system, B. Semantics for stronger systems can be obtained by adding various postulates as restrictions on the relations in the model structure [50, p. 300f]. Routley and Plumwood are not clear as to which extension of B to prefer. This may partly be a matter of context.

For the deontic system, Routley and Plumwood add a new, two-place relation, \(S\), where \(S(w, x)\) is read as “\(x\) is accepted from \(w\)”. The interpretation of the obligation operator is basically the same as in standard deontic logics: \(I(\Box a, w) = 1\) iff for every world, \(x\) such that \(S(w, x), I(a, x) = 1\).

We can now see how a moral dilemma of the form \(\Box(a \land \neg a)\) does not have to lead to any proposition’s being obligatory. Assume that we have just one accepted world, \(w\), (such that \(S(T, w)\)), and that \(a \land \neg a\) holds in \(w\). That \(\neg a\) holds in \(w\) just means that \(a\) does not hold in the “reversal” world \(w^\ast\). From this, it does not follow that any \(b\) holds in \(w\). We can also see why they have to use the complicated three-place R-relation instead of the ordinary accessibility relation in modal logic. Assume that we want to draw information out of a theory containing some impossibility, \(\neg(a \rightarrow a)\). For this to be true in a world \(w\), \(a \rightarrow a\) must not hold in \(w^\ast\). But then, there must, according to the interpretation of \(\rightarrow\), be two worlds \(x\) and \(y\) such that \(R(w, x, y)\) and \(a\) hold in \(x\) but not in \(y\).
Routley and Plumwood stress that the choice in dilemma situations often is not completely arbitrary. They have a special operator, $\otimes b$, which is to be read as “$b$ is the best thing, or an adequate thing, to do in the constrained circumstances” [49, p. 36], where $a$ and $\neg a$ are both forbidden. They have little to say about the semantics of this operator, but they do say that the best thing in a dilemmatic situation can be determined by “weighing up the expected outcomes and selecting among the relevant alternatives”, a roughly consequentialist decision procedure. They also point out that this need not be the same thing as a utilitarian weighing of the utilities of everybody affected.

6.6 Danielsson on Prescriptive and Ascriptive Norms

In his article “Preskriptiva och askriptiva normer” (“Prescriptive and Ascriptive Norms”), Sven Danielsson distinguishes between two kinds of norms. A “prescriptive” norm implies a corresponding imperative. For example, the normative expression “You ought to stop” implies, if interpreted as the expression of a prescriptive norm, the imperative “Stop” [18, p. 3].

On the other hand, “$p$ is wrong”, in the ascriptive sense, is said to be true iff $p$ is true and there is a prescriptive norm $\otimes (q)$ and a true $r$ such that $r \land p$, but not $r \land \neg p$, is inconsistent with $q$ [18, p. 3]. The important thing is that the prescriptive norms cannot give rise to dilemmas, but the ascriptive norms can, in some sense. Danielsson makes the following assumption of consistency between prescriptions and circumstances.

\[(CPC)\] The totality of what the prescriptive norms prescribes in a situation has to be consistent with the assumed circumstances, e.g. $\otimes (p_1) \land \ldots \land \otimes (p_n)$ holds in a situation only if the conjunction $p_1 \land \ldots \land p_n$ is compatible with the circumstances in that situation [18, p. 5].

But it may, for example, be that it would be ascriptively wrong to realize $p$, and also ascriptively wrong to realize $\neg p$, e.g. if we have a general norm which says that every promise ought to be kept, and have made incompatible promises. Every particular instance of the general norm can be fulfilled, but not without breaking another instance of the norm. We cannot derive a set of particular prescriptions that cannot be jointly satisfied (rather, the general norm fails to prescribe anything specific regarding the incompatible promises in the situation), but we can say that every alternative involves some break against an instance of the general norm [18, p. 7]. Danielsson thinks that most alleged examples of genuine dilemmas can be handled this way. In his article, he does not develop any complete formal system for his different kinds of norms.

If Danielsson’s approach to dilemmas is right, we do not have to make any changes to CDM1 in order to handle them. We can simply use the multiple
indexing to formalize situations where e.g. some instance of a general norm that holds relative to one index \( i \) has to be broken relative to another index \( j \); and we cannot, of course, say that the general norm holds relative to \( i \).

Let \( L \) denote a “promise-operator” so that \( L(p) \) stands for a proposition stating that there is a promise to \( p \), and let \( m \) stand for a proposition stating that someone is given money. We have the norms \( 
abla_i (L(m) \supset m) \land 
abla_i (L(\neg m) \supset \neg m) \). From this, we can derive the prohibitions \( F_i (L(m) \land \neg m) \land F_i (L(\neg m) \land m) \land F_i (L(m) \land L(\neg m)) \). Assume that we have another index, \( j \), where incompatible promises have been given, so that \( \Box_j (L(m) \land L(\neg m)) \) holds. One \( i \)-norm is already violated relative to \( j \), and some other violation, which will be realized by some \( j \)-open proposition, is \( j \)-necessary, because \( \Box_j [(L(m) \land \neg m) \lor (L(\neg m) \land m)] \land \Diamond_j (m) \land \Diamond_j (\neg m) \) holds.

Is this all we need in order to handle dilemmas? We may note that no norm that holds relative to the \( j \)-context is necessarily violated relative to that context. You do not violate any holding norm by realizing either \( m \) or \( \neg m \) in the \( j \)-context. But is it not characteristic of a dilemma situation that we have to violate one holding norm so that it is appropriate for us to feel guilty? In this type of situation, we might say that the feeling of guilt refers to the earlier violation we made in giving incompatible promises, whose negative consequences are actualized when we have to break one of the promises. But some alleged examples of dilemmas seem to be situations where someone is put in a tragic choice situation, through no fault of one’s own. One example would be a situation of the Sophie’s Choice type (see 1.5.5). Sophie is in a situation where she is forced to sacrifice one of her children, and she will presumably feel guilty whatever she does, but she has not been put in this situation through any earlier fault of her own, so, one might argue, if a norm, or set of norms, that forbids sacrificing any child do not hold in the situation where she has to make the sacrifice, she does not have anything to feel guilty about. On the other hand, I think most people evaluating Sophie’s actions from a third-part perspective would not consider her blameworthy for sacrificing any of her children, but rather pitiable, in contrast with the person who has given incompatible promises. It might then be plausible to say that Sophie never does anything wrong, but that she cannot help feeling guilty when she has to break one norm she has regarded as unconditionally valid. However, in section 6.7.1 below, I will discuss the possibility of assessing dilemmas with incompatible obligations in the CDM1 framework.

6.7 Incorporating a Solution into CDM

6.7.1 Dilemmas in the CDM Framework

One might try to make room for dilemmas in the NRDIM, and at the same time retain the validity of arguments like (Service), by removing the (DPos)
schema, which yields contradiction in case of dilemma. Let the system resulting from the removal of (DPos) from CDM1 be denoted CDM2.0.

CDM2.0 can handle dilemmas of the form $\Box_i(a) \land \Box_i(\neg a)$ (which implies $\Box_i(a \land \neg a)$, because of the (OAND) schema) without getting an inconsistency or the result that anything is obligatory. A dilemma of the Sophie’s Choice type, where we have an obligation whose fulfillment is not logically, but nomologically impossible, may be represented by a conjunction like $\Box_s \neg(c_1 \lor c_2) \land \Box_s (c_1 \lor c_2)$. It is easy to verify that arguments like (Service) are valid in CDM2.0, because of the (OAND) and (ORE) schemata.

It would appear natural to define a new operator in CDM2.0, in order to have something that fills the function of Goble’s obligation operator with permitted consequence principle.

(MRP) $\text{MRP}_i(a) \leftrightarrow \exists b(\Box_i(b) \land P_i(b) \land \Box_i(b \supset a))$, for any $i$

If $a$ is something that is $i$-obligatory, and there is no $i$-obligation that cannot be fulfilled without $\neg a$ being true, then anything that is $i$-necessitated by $a$ can be said to be a permitted requirement.

Any instance of this weaker schema is provable in CDM2.0.

(DPosP) For any $i$: $\vdash \text{MRP}_i(a) \supset \Diamond_i(a)$

$\Box_i(a \supset b)$ and $\Box_i(\neg b)$ together entail $\Box_i(\neg a)$ in ordinary modal logic. But $\Box_i(\neg a)$ and $\Box_i(a)$ entail $\neg P_i(a)$ in CDM2.0. Therefore, $\text{MRP}_i(a)$ is inconsistent with $\neg \Diamond_i(a)$.

However, we must remember that the relation between obligation and permission is different in the CDM systems, compared with Goble’s system, where the standard SDL definition $P(a) \leftrightarrow \neg \neg \Box \neg(a)$ is used, and this means that the MRP operator is not of much use in a context where we have a dilemma. If we have $\Box_i(a) \land \Box_i(\neg a)$ for any $a$ (which of course is implied by a statement like $\Box_i(b) \land \Box_i(\neg b)$ for any $b$), we also have $\text{MRP}_i(c)$ for any $c$. Nothing is sufficient for fulfilling all one’s obligations in such a situation. No statement of the form $P_i(d)$, and a fortiori no statement of the form $\text{MRP}_i(d)$, will thus be true. This is a consequence of the underlying structure in the CDM systems: we have traded a weaker concept of obligation for a stronger concept of permission, compared with SDL, in order to allow for a satisfactory handling of the Alf Ross Paradox and related problems, as explained in section 4.2.1.

But would not even people who believe e.g. that Sophie is faced with a genuine dilemma in her choice situation nevertheless believe that she is permitted to e.g. scratch her head or to sit down and cry in that situation? CDM2.0 rules this out, and is therefore unacceptable as an approach to handling moral dilemmas. I have to admit that I see no way to handle dilemmas of the type $\Box_i(a \land \neg a)$ in the CDM1 framework in a satisfactory way.

However, dilemmas of the type $\Box_i(a) \land \Box_i(\neg a)$ can be handled in the CDM framework if we replace the (OAND) schema with the weaker:
Let the system resulting from replacing (OAND) with (PosAND) in CDM1 be denoted CDM2.1. It is easy to see that arguments like (Service) are valid in CDM2.1, because we may assume that \( f \lor s \) and \( \neg f \) are jointly possible; but we cannot derive \( i (a \land \neg a) \) from \( i (a) \land i (\neg a) \), and we cannot, in the Sophie situation, derive \( i (\neg c1 \land \neg c2) \) from \( i (\neg c1) \land i (\neg c2) \), because we have assumed \( i (c1 \lor c2) \). Generally, we can say that we have a dilemma in relation to the index \( i \) if it holds that \( \exists a \exists b \circ_i (a) \land \circ_i (b) \land \neg \circ_i (a \land b) \), or, equivalently, \( \exists a \exists b \circ_i (a) \land \circ_i (b) \land \neg \circ_i (a \land b) \).

It is plausible that the prohibition that Sophie sacrifices any child comes from a general norm against sacrificing one’s children. This could be expressed in a language with quantification over ordinary individuals, but we must be careful with the formalization. Let \( P(x, y) \) stand for “\( x \) is a parent of \( y \)” and \( S(x, y) \) stand for “\( x \) performs an act of sacrificing \( y \)”.

This type of solution does not work for an obligation operator with consequence principle [29, p. 468]. If we have assumed \( \circ a \land \circ \neg a \), we can then, for any \( b \), derive \( \circ (a \lor b) \), and, if \( b \) and \( \neg a \) are consistent (Goble speaks of consistency rather than possibility), then also derive \( \circ ((a \lor b) \land \neg a) \), which is equivalent to \( \circ b \). If \( b \) is consistent with \( a \) but not \( \neg a \), we can, of course, make a similar derivation. So, since any consistent proposition must be consistent with either \( a \) or \( \neg a \), we have the absurd result that \( \circ b \) holds for any consistent \( b \) if there is a dilemma of the form \( \circ a \land \circ \neg a \).

But this is not a problem in CDM2.1. The consequence principle does not hold for the \( \circ \)-operators, so we cannot derive \( \circ_i (a \lor b) \) from \( \circ_i a \). If we have \( \circ_i a \land \circ_i \neg a \), we can indeed derive \( \text{MR}_i (a \lor b) \land \text{MR}_i (\neg a) \) for any \( b \), but these requirements cannot be joined to \( \text{MR}_i b \), because no principle of aggregation holds for the MR-operators \(^2\).

On the semantic level, we can accommodate the restriction on (OAND) by stating that \( wS_i X \) and \( wS_i Y \) imply \( wS_i X \cap Y \) provided that there is some world \( w' \) such that both \( wR_j w' \) and \( w' \in X \cap Y \).

We may note that there is no difference between CDM1 and CDM2.1 in cases where we have no conflicting obligations, so the CDM1 assessment of the SDL problems described in chapter 4 remains valid in CDM2.1.

---

\(^2\)In a non-dilemmatic case, where \( \text{MR}_i a \land \text{MR}_i b \) holds because some jointly possible \( c \) and \( d \) are both obligatory, \( \text{MR}_i (a \land b) \) may be derived from \( \circ_i (c \land d) \).
6.7.2 A DWE-like Assessment of Supererogation

Among the different approaches to handling supererogation in deontic logic discussed above, McNamara’s DWE comes closest to the basic ideas of CDM1, because it contains no reference to agency or values, and I cannot see that it has any serious philosophical disadvantages compared with Chisholm’s proposals. I propose that we incorporate the main ideas behind McNamara’s DWE into a CDM1-like system in the following way. In the system CDM3, we add the following elements to CDM1 (cf. section 2.4):

1. Indexed versions of McNamara’s primitive operators: MI$_i(a)$, MA$_i(a)$ and IN$_i(a)$.
2. All instances of these axiom schemata, the analogue of McNamara’s (A1) schema: (MIK) MI$_i(a \supset b) \supset (MI_i(a) \supset MI_i(b))$ and (MAK) MA$_i(a \supset b) \supset (MA_i(a) \supset MA_i(b))$
3. All instances of the analogue of McNamara’s (A2) schema: (CA2) MR$_i(a) \supset (MI_i(a) \land MA_i(a))$
4. The (DPos) schema is replaced by the analogue of McNamara’s (A3) schema: (CA3) MI$_i(a) \lor MA_i(a) \supset \circ_i(a)$
5. All instances of the analogue of McNamara’s (A5) schema: (CA5) IN$_i(a) \supset P_i(a)$
6. All instances of the analogue of McNamara’s (A6) schema: (CA6) IN$_i(\neg a)$
7. All instances of the analogue of McNamara’s (A7) schema: (CA7) IN$_i(a) \supset \neg (MI_i(a) \lor MA_i(a))$
8. All instances of the analogue of McNamara’s (A8) schema: (CA8) [MR$_i(a \supset b) \land MR_i(\neg a \supset \neg c) \land IN_i(b) \land IN_i(c)] \supset IN_i(a)$
9. All instances of the necessitation schema for the MA and MI operators: (NecD) $\Box_i(a) \supset MI_i(a) \land MA_i(a)$

I propose the following definitions analogous to McNamara’s derived operators.

1. GR$_i(a) \leftrightarrow_{def} \neg MR_i(a)$
2. OP$_i(a) \leftrightarrow_{def} P_i(a) \land \neg MR_i(a)$
3. SI$_i(a) \leftrightarrow_{def} \neg IN_i(a)$
4. SU$_i(a) \leftrightarrow_{def} P_i(a) \land MI_i(\neg a)$
5. SO$_i(a) \leftrightarrow_{def} P_i(a) \land MA_i(\neg a)$

I have thus followed McNamara in treating MI and MA as concepts of deontic necessity, like the obligation operator in SDL. I have also defined the analogues of McNamara’s derived operators, except for the moral significance operator, in terms of the CDM3 derived operators of moral necessity, and not the primary operators. I think these corresponding natural language expressions are closely related to necessary conditions for different types of duty fulfillment. For example, if $a$ is something obligatory, we would not say that $a \lor b$ is something optional or gratuitous, and we should thus welcome that $\Box_i(a) \supset \neg GR_i(a \lor b)$ is a theorem in CDM3, which is easily verified. Some-
thing that is not forbidden but excludes doing one’s best, should, if we believe that this deontic category is possible at all, be called suboptimal. The supererogatory is that which is permitted but excluded by only fulfilling the minimal requirements.

I propose that we represent CDM3 by the CDM3S semantics, where the following elements are modified or added in relation to the CDM1S semantics:

1. A CDM3S model is an ordered septuple $<W, R, S, T, P, V, \le>,$ with the elements $W, R, P$ as in an CDM1S model.

2. $T$ is a set of binary relations $T_i, T_j, T_k \ldots$ between worlds, like $R_i$. They should be thought of as more traditional SDL-like deontic accessibility relations. For any index $i$, and any worlds $w, w', wT_iw'$ holds iff $wR_iw'$ holds and $w' \in \{\cap X | wS_i X\}$.

3. $\le$ is a set of tenary relations $\le_i, \le_j, \le_k \ldots$, where the informal interpretation of $w_1 \le_i w w_2$ is that world $w_2$ is ranked at least as high as $w_1$, relative to index $i$ and world $w$. An $\le_{iw}$-ranking forms a weak order, i.e. is transitive, so that for any $i, w, w_1, w_2, w_3, w_1 \le_{iw} w_2$ and $w_2 \le_{iw} w_3$ implies $w_1 \le_{iw} w_3$, and total, so that for any $w_1$ and $w_2$ such that $wT_iw_1$ and $wT_iw_2$ holds, either $w_1 \le_{iw} w_2$ or $w_2 \le_{iw} w_1$ holds.

4. The $V$ valuation functions as in an CDM1S model, except for the truth-conditions for the new operators. The set $|a|$ corresponding to a proposition $a$ is defined as for CDM1S models.

5. For any world $w$, and any MA-operator $MA_i$, if there is a world $w_1$ such that $wT_iw_1$, and it holds for any world $w_2$ that if $w_1 \le_{iw} w_2$, then $V(a, w_2 = 1)$, then $V(MA_i(a), w = 1)$, otherwise $V(MA_i(a), w = 0)$.

6. For any world $w$, and any MI-operator $MI_i$, if there is a world $w_1$ such that $wT_iw_1$, and it holds for any world $w_2$ that if $w_2 \le_{iw} w_1$, then $V(a, w_2 = 1)$, then $V(MI_i(a), w = 1)$, otherwise $V(MI_i(a), w = 0)$.

7. For any world $w$, and any IN-operator $IN_i$, if it holds for any $w_1$ such that $wT_iw_1$, there is a $w_2$ such that $w_2 =_{iw} w_1$ and $V(a, w_2 = 1)$, and also a $w_3$ such that $w_3 =_{iw} w_1$ and $V(\neg a, w_3 = 1)$, then $V(IN_i(a), w = 1)$, otherwise $V(IN_i(a), w = 0)$.

I have just adapted McNamara’s conditions to the CDM framework and defined his new operators in relation to the MR-operators, that function like standard deontic necessity operators like the $\bigcirc$-operators in McNamara’s system. Therefore, adapted versions of McNamara’s proofs of soundness and completeness in conjunction with the proofs given in 2.4.4 should be sufficient to show that CDM3 is represented by CDM3S.

If we want to allow for both dilemmas and supererogation in one system, I propose that we use the system CDM4, where the CDM3 modifications are made to CDM2.1 instead of CDM1. We cannot interpret the (CA2) and (CA8) schemata and the $T_i$-relation exactly as in CDM3, because, in the case of a dilemma relative to a world $w$ and an index $i$, the set of $i$-accessible worlds where all obligations are satisfied will be empty, and the system would collapse – $MI_i a$ could be derived for any $a$, which would contradict the modified
Figure 6.1: Illustration of the CDM3 semantics. Proposition $a$ is supererogatory, while proposition $b$ is obligatory in world $w$ relative to index $i$. $R_i$ is a weak modal accessibility relation, and $S_i$ is a deontic accessibility relation. For all worlds $w'$ that are deontically accessible relative to $i$ and $w$, i.e. for which $wT_iw'$ holds, all $\neg a$-worlds are ranked lower than all $a$-worlds. Cf. the CDM1 semantics illustrated in figure 2.1.

(CA3) schema. One solution would be to only let moral requirements that do not come into conflict with other requirements define the set of acceptable states, i.e. to use MRP instead of MR in the (CA2) and (CA8) schemata. We should then also modify the second condition in the CDM3S semantics so that $wT_iw'$ holds iff $wR_iw'$ holds and $w' \in \cap X|wS_iX$ and there is no $Y$ such that $wS_iY$ and $X \cap Y = \emptyset$.
7. Appendix

In this appendix, I list axioms, definitions, rules, and some important theorems in the systems SDL and CDM1, because these are referred to throughout the dissertation. I will also give a summary comparison as to how the different systems discussed in the dissertation fare with respect to the conditions stated in 1.9.

7.1 SDL Formulae

Definitions:
1. \( P(a) \leftrightarrow \neg \bigcirc (\neg a) \)
2. \( F(a) \leftrightarrow \bigcirc (\neg a) \)

Rules and axiom schemata:

(K) \( \vdash \bigcirc (a \supset b) \supset (\bigcirc (a) \supset \bigcirc (b)) \)

(D) \( \vdash \bigcirc (a) \supset P(a) \)

(RN)

\[
\begin{align*}
\vdash a \\
\vdash \bigcirc (a)
\end{align*}
\]

Equivalently:

(RE)

\[
\begin{align*}
\vdash a \leftrightarrow b \\
\vdash \bigcirc (a) \leftrightarrow \bigcirc (b)
\end{align*}
\]

(D) \( \vdash \bigcirc (a) \supset P(a) \)

(\&E_d) \( \vdash \bigcirc (a \land b) \supset \bigcirc (a) \)

(\&I_d) \( \vdash \bigcirc (a) \land \bigcirc (b) \supset \bigcirc (a \land b) \)

(EM_d) \( \vdash \bigcirc (a \lor \neg a) \)
Derived rules and theorems:

(RR) (Also known as “the consequence principle”, or “the rule of inheritance”.)

\[
\begin{align*}
\vdash a & \supset b \\
\vdash \lozenge(a) & \supset \lozenge(b)
\end{align*}
\]

\(\lor I_d\) \(\lozenge_i(a) \supset \lozenge_i(a \lor b)\)

7.2 CDM1 Formulae

Definitions:

1. \(\lozenge_i a \leftrightarrow_{\text{def}} \neg \square_i \neg a\), for any \(i\).
2. \(\text{MR}_i a \leftrightarrow_{\text{def}} \exists b(\lozenge_i(b) \land \square_i(b \supset a))\), for any \(i\).
3. \(P_i a \leftrightarrow_{\text{def}} \neg \text{MR}_i \neg a\), for any \(i\).
4. \(F_i a \leftrightarrow_{\text{def}} \text{MR}_i \neg a\), for any \(i\).
5. \(\text{MRP}_i a \leftrightarrow_{\text{def}} \exists b(\lozenge_i(b) \land P_i(b) \land \square_i(b \supset a))\), for any \(i\).
6. \(\text{MS}_i(a) \leftrightarrow_{\text{def}} P_i(a) \land \forall b(\lozenge_i(b) \supset \square_i(a \supset b))\)
7. \(\text{MSR}_i(a) \leftrightarrow_{\text{def}} \text{MS}_i(a) \land \text{MR}_i(a)\)

Rules and axiom schemata:

(K) For any \(i\): \(\vdash \square_i(a \supset b) \supset (\square_i(a) \supset \square_i(b))\)

(Nec) For any \(i\):

\[
\begin{align*}
\vdash a \\
\vdash \square_i(a)
\end{align*}
\]

(T) For any \(i\): \(\vdash \square_i(a) \supset (a)\)

(ORE) \(\vdash [\lozenge_i(a) \land \square_i(a \leftrightarrow b)] \supset \lozenge_i(b)\)

(OAND) For any \(i\) \(\vdash \lozenge_i(a) \land \lozenge_i(b) \supset \lozenge_i(a \land b)\)

(OOR) For any \(i\) \(\vdash \lozenge_i(a) \land \lozenge_i(b) \supset \lozenge_i(a \lor b)\)

(DPos) For any \(i\) \(\vdash \text{MR}_i(a) \supset \lozenge_i(a)\)

(E\forall) \(\vdash \forall p(a) \supset a[q/p]\), where \(q\) is free for \(p\) in \(a\).

(K\forall) \(\vdash \forall p(a \supset b) \supset (\forall p(a) \supset \forall p(b))\)

(N\forall) \(\vdash a \supset \forall p(a)\), where \(p\) is not free in \(a\).

(BF) For any \(i\) \(\vdash \forall p \square_i(a) \supset \square_i \forall p(a)\)
(G) \[ \vdash a \quad \vdash \forall p(a) \]

(DBF) In addition to Fine’s proof system, Barcan Formula also for the obligation operators: For any \( i \): \( \vdash \forall p \diamond_i (a) \supset \diamond_i \forall p(a) \)

Theorems:

(DPosP) For any \( i \): \( \vdash \mathrm{MRP}_i(a) \supset \diamond_i(a) \)

(PPos) For any \( i \): \( \vdash \diamond_i(a) \supset [P_i(b) \supset \diamond_i(b)] \)

(OMSR) For any \( i \): \( \vdash \mathrm{MSR}_i(a) \supset \diamond_i(a) \)
7.3 Comparison Table for Systems

<table>
<thead>
<tr>
<th>System</th>
<th>Ref</th>
<th>ROI</th>
<th>NROI</th>
<th>C</th>
<th>P</th>
<th>MC</th>
<th>SE</th>
<th>Dil</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDL</td>
<td>1.4</td>
<td>–</td>
<td>–</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CDM1</td>
<td>2</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>D_0**</td>
<td>3.2</td>
<td>–</td>
<td>–</td>
<td>x</td>
<td>–</td>
<td>x</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>MO/3-D</td>
<td>3.3</td>
<td>x</td>
<td>–</td>
<td>x</td>
<td>x</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>DWE</td>
<td>6.3</td>
<td>–</td>
<td>–</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>–</td>
</tr>
<tr>
<td>DPM</td>
<td>6.4.2</td>
<td>–</td>
<td>–</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>–</td>
<td>x</td>
</tr>
<tr>
<td>(Routley/Plumwood)</td>
<td>6.5</td>
<td>–</td>
<td>–</td>
<td>x</td>
<td>–</td>
<td>–</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>CDM2.1</td>
<td>6.7.1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>–</td>
<td>x</td>
</tr>
<tr>
<td>CDM3</td>
<td>6.7.2</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>–</td>
</tr>
<tr>
<td>CDM4</td>
<td>6.7.2</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 7.1: This is a summary comparison as to how the different systems discussed in the dissertation fare with respect to the conditions stated in 1.9. Index: ROI: the system contains a possibility-indexed moral requirement operator for which (RR) or a similar principle of regularity holds, NROI: the system contains a possibility-indexed obligation operator for which (RR) or a similar principle of regularity does not hold, C: the system is built on classic logic, P: the deontic operators operate on propositions, MC: the system uses the material conditional for conditional obligations, SE: the system handles supererogation, Dil: the system handles genuine obligation dilemmas, where we have two jointly unsatisfiable states that are both obligatory. I have argued that a deontic system that is adequate for the purposes defined in section 1.3 must satisfy both ROI and NROI. I think it is desirable that it satisfies C–MC as well, because of considerations of theoretical simplicity. The names of systems defined in this dissertation are printed in bold. I leave it open whether a system that fulfills the ultimate goal of making a fully adequate translation of natural deontic reasoning into the formal language of the system possible (as stated in section 1.3) also should satisfy SE or Dil.
Acknowledgments

I would first like to thank my main supervisor, Erik Carlson, my assistant supervisor during the first two years of work, Jan Österberg, and my assistant supervisor during the last two years, Kent Hurtig. Sven Danielsson also deserves special mention, because his ideas about deontic logic, especially regarding the Alf Ross Paradox and the interdefinibility of the deontic operators, have served as a great inspiration in the construction of the CDM systems, and so does John Cantwell, who acted as opponent on a draft of the whole dissertation at the final seminar.

Then, I would like to thank the other participants at the Higher Seminar in Practical Philosophy at Uppsala University who have commented on drafts of the different chapters: Per Algander, Thomas Anderberg, Karin Enflo, Per Ericson, Traugott Schiebe, the late Howard Sobel, Torben Spaak, Frans Svensson, and Folke Tersman, as well as various participants at the Graduate Seminar in Practical Philosophy at Stockholm University, who have commented on chapter drafts discussed there: Katharina Berndt, Johan Gustafsson, Sofia Jeppsson, Jens Johansson, Jonas Olson, Niklas Olsson-Yaouzis, and Torbjörn Tännsjö, the participants at the Formal Ethics Week in Groningen, September 2010, especially Davide Grossi and Barteld Kooi, and other people who have given comments, including Daniel Rönneval and Kim Solin.

Also, thanks to David McVicker, for checking my English, and to Lars-Göran Johansson and Rysiek Sliwinski for help with various practical matters.


[34] Ingemar Hedenius. *Om rätt och moral*. Tidens, 1941.


Åqvist, L., 6, 65
B (axiom), 11
Belzer, M., 42
BF (axiom), 34, 88
C (adequacy condition), 22, 90
Carlson, E., 26
Castañeda, H.-N., 5, 9, 46, 64
CDM1, 5
CDM2 (CDM adaptation for dilemmas), 80–82
CDM3 (CDM adaptation for supererogation), 83–84
CDM4 (CDM adaptation for dilemmas and supererogation), 84–85
Chisholm, R., 69
D (axiom), 9, 87
Dahlquist, T., 13
Danielsson, S., 26
Dil (adequacy condition), 22, 90
dilemmas
rejection of the (K) schema, 75–76
relevant logic, 77–79
restricting the consequence principle, 76–77
DPM, 76–77
DPs (axiom), 34, 88
DPsP (theorem), 81, 89
Dr 11.1 (rule), 53
DWE, 73–75
E (axiom), 34, 88
Eck, J.A. van, 64, 66
\( \wedge E_d \) (axiom), 10, 87
\( \wedge EM_d \) (axiom), 10, 87
expressivism, 18
Fieldman, F., 43
Fine, K., 24
first-order logic, 8
Forrester, J.W., 15, 66
G (rule), 34, 89
Geach, P., 60
Goble, L., 76
Gustafsson, J., 24, 74
Hansen, J., 47
Hansson, B., 19
Hansson, S.O., 71
Harman, G., 15
Hedenius, I., 19
specification, 41
deontic categories
in terms of dyadic requirement, 72
Diladequacy condition, 22, 90
dilemmas
and Good Samaritan Paradoxes, 49–50
formal specification, 33
informal specification, 23–32
necessity, 29
obligation vs. moral requirement, 24
CDM2 (CDM adaption for dilemmas), 80–82
CDM3 (CDM adaption for supererogation), 83–84
CDM4 (CDM adaption for dilemmas and supererogation), 84–85
Chisholm, R., 69
D (axiom), 9, 87
Dahlquist, T., 13
Danielsson, S., 26
on Alf Ross’s paradox, 46–48
on Chisholm’s Paradox, 59–60
on dilemmas, 79–80
DBF (axiom), 34, 89
D** (Castañeda’s systems)
and Chisholm’s Paradox, 56–57
and Good Samaritan Paradoxes, 52–53
Hansen, J., 47
Hansson, B., 19
Hansson, S.O., 71
Harman, G., 15
Hedenius, I., 19
Horty, J., 15, 70
Hume’s Law, 20
∧I_d (axiom), 10, 87
Jackson, F., 55
Jennings, R.E., 75
K (axiom), 9, 34, 87, 88
KV (axiom), 34, 88
Kanger, S., 63
Key (rule), 43
Lewis, D., 15, 42
Lewis-style systems
3-D, 43
and Chisholm’s Paradox, 57–59
and Good Samaritan Paradoxes, 53
MO, 43
V-logics, 42
Loewer, B., 42
Mackie, J.L., 28
MC (adequacy condition), 22, 90
McNamara, P., 73
minimalism (theory of truth), 18
multiplied beneficence argument, 66–68
NV (axiom), 34, 88
Nec (rule), 88
NROI (adequacy condition), 22, 90
OAND (axiom), 34, 88
OMSR (theorem), 29, 89
OOR (axiom), 34, 88
ORE (axiom), 34, 88
P (adequacy condition), 22, 90
P* (principle), 53
Parfit, D., 32
Pargetter, R., 55
Pigden, C., 21
Plumwood, V., 69, 77
PosAND (axiom), 82
PPos (theorem), 28, 89
P’ (principle), 52
practitions, 39–41
Prior, A.N., 12, 20
quantified deontic logic
Jephta Dilemma, 65
need, 17–18
relation to modal logic, 63–64
substitution of individuals, 65
RE (rule), 10, 87
RN (rule), 10, 34, 87
ROI (adequacy condition), 22, 90
Routley, R., 69, 77
RR (rule), 10, 88
S4 (axiom), 11
Salwén, H., 21
Schurz, G., 20, 63, 64
Scotch, P.K., 75
SDL, 6
problems
Alf Ross’s Paradox, 12, 45–48
Chisholm’s Paradox, 13–14, 54–60
conditional obligations, 12–14
dilemmas, 15
Good Samaritan Paradox, 12, 49–53
John/Tom-example, 14, 60–61
supererogation, 16
scope solution, 50–52
specification, 9–10
SE (adequacy condition), 22, 90
Segerberg, K., 23
Singer, P., 67
Smith, H., 55
Sosa, E., 71
supererogation
in terms of dyadic requirement, 72
in terms of evaluation of actions, 69–72
in terms of McNamara’s operators, 73–75

T (axiom), 11, 34, 88
Tännö, T., 19

Wedberg, A., 51
Wright, G.H. von, 9, 11, 42, 65