

Igniting the Deontic Consequence Relation: *Dilemmas, Trumping, and the Naturalistic Fallacy*

By

Kurt Holukoff

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Arts
in
Philosophy

Waterloo, Ontario, Canada, 2007

©Kurt Holukoff 2007

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this work, Kurt Holukoff examines three formal approaches to representing valid inferences in reasoning regarding obligation and its cognates: deontic logic. He argues that an appropriate formalization of deontic logic should take genuine moral dilemmas seriously, be capable of representing trumping-like reasoning, and not make the naturalistic fallacy valid as a matter of logic. The three systems he investigates are, the Standard Deontic logic, a Relevant Deontic logic, and Schotch and Jennings' multiple moral accessibility relations Deontic logic. The Standard Deontic logic has seemingly insurmountable problems representing both fruitful reasoning from an inconsistent set of obligations and trumping-like reasoning. Moreover, the naturalistic fallacy is valid in the Standard Deontic logic. The Relevant deontic logic that the author examines is capable of representing fruitful reasoning from an inconsistent set of obligations and does not make valid the naturalistic fallacy. However, the author argues that the Relevant deontic logic needs some revisions in order to represent trumping-like reasoning. Likewise, the author finds that Schotch and Jennings' Deontic logic is capable of representing fruitful reasoning from an inconsistent set of obligations. However, in order to represent trumping-like reasoning, revisions to Schotch and Jennings' Deontic logic are apparently required. Similar revisions are seemingly required to block the naturalistic fallacy, which is otherwise valid in Schotch and Jennings' original system.

Acknowledgements

I was very happy to write this thesis surrounded and supported by good friends, all of them excellent philosophers. In particular, Eric Hochstein read early drafts and his advice went far in keeping me on track. My wife, Robyn, put her wonderful editing skills to work, without which, I am certain that I would have been hard pressed to finish this work in any reasonable amount of time. It is also due to Robyn's insistence that I tried to constantly provide a philosophical interpretation to the different semantics I was using. Dr. Dave DeVidi's constant encouragement, guidance and advice were invaluable. Any competence I hopefully developed in writing philosophical logic during this experience is due in large part to him.

Table of Contents

Chapter One: Standard Deontic logic.....	1
Chapter Two: Relevant Deontic logic.....	25
Chapter Three: Schotch and Jennings' Deontic logic.....	43
Chapter Four: Directions for Further Research.....	54
Appendix A: Truth conditions for Trumping.....	58
Appendix B: A Negation for Deontic logic.....	61

Chapter One

Obligation as Moral Necessity: in Classical Terms

In 2004, Jillian Searle¹ and her two young children were enjoying breakfast on the Thai island of Phuket. However, within the span of a few moments, the tsunami that devastated Indonesia and Thailand forced Searle into make a horrifying decision. Searle realized that with the waters rising and threatening to sweep her two young children out to sea, she did not have the ability to hold on to both children. If she did not let go of one child, all three of them would perish.

Luckily for Searle, one of her children was five years old and she thought that he at least had a slight chance of surviving on his own, whereas the other child was still a baby. Searle let go of the older child. Although he was nearly swept out to sea, fortunately he was able to hold onto a door to keep his head above water. Searle and her baby also survived the tsunami. While the decision that Searle was forced to make was indeed horrific, it could have been worse if there were no relevant considerations to help her select one child over the other. Perhaps they could have been twins, or both toddlers. It seems that if the decision had to be made under those conditions it should be considered an impossible moral situation or alternatively, a genuine moral dilemma.

Consider then, a reasonable person, with a normal range of emotions and sympathies, who has discovered that such genuine moral dilemmas can occur and have occurred. Since, at least initially, genuine moral dilemmas may be explicated as inconsistent sets of moral demands, in all of her future moral inquiries, if she follows the norms of appropriate reasoning as taught her in her Logic 101 course, she is entitled to infer that anything and nothing is morally obligatory and permissible. For *Explosion*, by

¹ From BBC website, <http://news.bbc.co.uk/2/hi/asia-pacific/4137053.stm>, accessed on August 20, 2007

which we mean the rule, that from an inconsistent set of premises anything follows, is a valid inference in classical logic and its modal affiliates. However, any type of moral inquiry which actually results in justifying such conclusions, i.e., that anything and nothing is morally obligatory and permissible, would likely be anathema to any person who believes that moral demands are meaningful and in some sense binding on people like her.

At this point, if we want to avoid such a disastrous account of moral inquiry, it seems that we have, broadly speaking, two options. First, we may deny the possibility of genuine moral dilemmas. Secondly, we may reject the validity of *explosion* in moral inquiry. The first option seems to run counter to the sad facts of moral experience: people have been, and unfortunately will likely continue to be, faced with unavoidable decisions under conditions of incompatible moral demands. The main result of this current project is that it is the second option, that is, by rejecting *explosion* as a rule of inference in moral inquiry, which permits us to make progress in formalizing reasoning regarding obligation and its cognates.

However, it is currently more common for moral theorists to reject the possibility of genuine moral dilemmas. In fact, it is standard practice when constructing informal ethical systems, for philosophers to try to provide theoretical mechanisms to reconcile all apparent moral conflicts; and these mechanisms seemingly entail that any putative moral dilemma is in fact resolvable and thus not a genuine moral dilemma at all.

Therefore, many philosophers seemingly hold that an essential element of a ‘good’ ethical theory is that it can show that there is *always* one obligation which can *trump* or over-ride other obligations when a conflict seemingly arises. For some, a good

informal theory ought to provide a rigid hierarchy of obligations; other theorists argue that good theories ought to provide a moral calculus; and some may believe that the use of ‘intuition pumps’ and thought experiments, even if in a somewhat *ad hoc* manner, will result in the best model of how one obligation should trump another obligation.

These methods are understood by their respective proponents as a necessary feature of their theories, in virtue of the requirement that informal moral theories be action-guiding. This ‘action-guiding’ criterion is often justified by the claim that it is this very feature which makes a theory a viable candidate for a good moral theory: it must always let an agent know which obligation is genuinely binding over other seemingly conflicting obligations. It is often said, and correctly so, that ‘action-guiding’ and ‘taking genuine moral dilemmas seriously’ are incompatible criteria for any informal moral theory.

It is the incompatibility of ‘being action-guiding’ in this very strong sense and ‘taking genuine moral dilemmas seriously’, together with the above-mentioned fact of moral experience (that there are genuine moral dilemmas) which lead me to deduce that there are limits to the guiding power of moral reasoning and theories. Thus, I propose that the present work be understood in part as a critique of moral reasoning: we will discover and explicate the limits of the deliverances of moral judgment, inquiry and theories.

The Standard Deontic Logic

The purpose of this chapter is to demonstrate that the most commonly accepted logic used to formalize moral reasoning, Standard Deontic Logic (SDL), does not deserve to be considered a good formalization. In what follows, I will offer a sketch of SDL and its deficiencies as a formalization of moral reasoning. The main reasons for rejecting

SDL is its inability to express statements regarding reasoning from an incompatible set of obligations and the trouble SDL has in expressing how *trumping* reasoning can work at all. A further reason to reject SDL is that it validates the deduction of an ‘ought’ from an ‘is’. I propose that it is likely inappropriate to use a formal system to characterize valid moral reasoning, if that system contains a valid formula which may be interpreted as justifying the naturalistic fallacy.

It seems initially plausible to equate moral obligation with quantifying over morally accessible possible worlds: obligation is understood as moral necessity as it were. When we say that a possible world is morally accessible from another world we mean that the possible world is *morally ideal* from the perspective of the world to which it is accessible. Thus, we may infer from the recognition of an obligation in our world, that in every world morally accessible from ours, the course of action prescribed by the obligation is true of that world². The inference is correct from right to left as well. Thus, given that some proposition holds in every relevant world morally accessible from ours, we may infer that there is an obligation to make that proposition obtain in our world.

Moreover, many theorists propose that, given an obligation, there must always be at least one relevant world morally accessible from ours. Therefore, not only is there at least one morally relevant possible world in which all obligations in our world are there satisfied, but in every morally relevant possible world, every obligation we could incur in our world, is in fact there satisfied. Given the traditional notion that ‘ought’ implies ‘can’, where ‘can’ is understood as requiring more than mere logical or physical possibility, morally relevant possible worlds must be realistically accessible through human means. Thus not only is there a possible world in which all our obligations are fulfilled, but

² I.e. is true of us, or our counterpart, in that world etc

according to traditional notions of morality, this possible world is similar enough to ours that we could transform our world into it through human means.

In order to provide a heuristic device as an aid in understanding why the checking of other worlds may help us in analyzing obligation in our world, consider the following: Suppose that there is a set of worlds, that includes all worlds that are morally ideal or desirable relative to our actual world, but where each member of this set is still sufficiently similar to our world that it is ‘achievable’ in a suitable sense.

Thus if something is true of at least one of these worlds, then it seems that it must at least be morally permissible for us to make it true in our world. For obligation, suppose that there is no world, morally ideal but still sufficiently similar to our world, in which **P** holds. Then we would not be permitted to make **P** true, and if **P** is impermissible than $\sim\mathbf{P}$ is obligatory³.

The use of possible world semantics, as in the above heuristic device, has led some theorists to investigate the applicability of a Kripke-style semantics to moral reasoning. To introduce the standard system of representing moral reasoning, SDL, we need the following definition of a *frame* in a Kripke-style semantics:

Definition 1.0 For our current purposes, a *frame* is a pair $\langle W, R \rangle$, where W is a non-empty set of worlds and R is a dyadic accessibility relation on W . So if xRy , we interpret this as saying that world y is accessible from world x .

This notion of a *frame* may then be used to sketch out what sort of formulas we would want to be valid for moral reasoning.

³ This heuristic device relies on the law of excluded middle and De Morgan negation; if this is problematic, imagine that **P** holds in every morally desirable world sufficiently similar to the actual world. An argument by analogy may then indicate that a moral agent must make **P** true of this world on pain of inconsistency.

As noted above, real world obligation is characterized by many theorists as implying that in every world morally accessible from ours, the course of action prescribed by the obligation is true of that world. Thus, if in our world (call it world- x for now) there is an obligation to do **A**, then for any world y , if xRy , **A** holds at that world y . Moreover, the inference holds for the argument right to left as well. We will use the obvious notation $\Box P$ (at x) to represent: P obtains at any w , if xRw ; and $\Diamond P$ (at x) to represent: there is at least one w such that xRw and that P holds at w .

Also, many theorists believe as a conceptual truth, that if anything is obligatory, then it must also be permissible. Recall also that **P** is permissible if, and only if, **P** is the case in at least one world morally accessible from ours; thus if there is an obligation in our world, there must always be at least one relevant world morally accessible from ours.

Thus the relation R is understood as being *serial*:

Definition 1.1 A relation R is serial iff for any w , there is an x , such that wRx .

This means that if at x there is an obligation to do **A**, one is always entitled to infer that there is at least one world morally accessible from x , that is some world y , at which **A** holds. Thus, given the above characterization, if $\Box P$ at x then $\Diamond P$ at x also.

Consequently, it seems that those who would find the above explication of moral reasoning satisfying thus far may also be committed to arguing that it is only those formulas valid in *serial frames* which have a chance at being parts of a good formalization of moral reasoning. To make the converse inference—that all formulas valid in such frames are parts of such a formalization—is probably somewhat hasty. As noted earlier, many philosophers believe that ‘ought’ not only implies ‘permissible’ but also ‘can’.

Thus, we may need more than one accessibility relation R for our Kripke *frames*, in order to represent both alethic necessity and possibility and moral obligation and permissibility within the same set W of worlds. Letting R° represent the *serial* relation, and R^2 represent the alethic relation, we assume (as is usually done) that R^2 is *reflexive*, *transitive*, and *symmetric*.

Definition 1.2 A relation R on W is *reflexive*, when for any w , wRw ;

Definition 1.3 A relation R on W is *transitive*, when for any x, y, z if xRy & yRz then xRz .

Definition 1.4 A relation R on W is *symmetric* when for any w, x if wRx then xRw .

R^2 is none other than the accessibility relation for the class of *frames* for the familiar modal logic **S5**.

The normal modal logic **KD** is the class of formulas valid in *serial frames*. But since SDL also validates the inference from ‘ought’ to ‘can’, **KD** is not sufficient. Therefore, philosophers who equate moral obligation with moral necessity (classically understood) may find a polymodal logic with **KD** and **S5** as a good candidate for formalizing moral reasoning; and indeed many have formalized SDL with the axioms of a polymodal logic combining **KD** and **S5**. Thus it must seem, at least initially, plausible that by using an appropriate interpretation, one may then use **KD** together with **S5** to express or prove any sentence involved in reasoning about *obligation* and its cognates.

KD is the smallest modal logic which contains every instance of the schemas

$$\Box(P \rightarrow Q) \rightarrow (\Box P \rightarrow \Box Q) \quad (\text{K})$$

$$\Box P \rightarrow \Diamond P \quad (\text{D})$$

That is, with a language of classical propositional logic with the standard set of propositional variables, the operators \sim, \rightarrow , to which we add two new unary operators, \Box , and \Diamond , **KD** may be axiomatized thusly:

A1. All classical tautologies of the language are axioms.

For any sentences P and Q , the following are valid:

A2. $\Box(P \rightarrow Q) \rightarrow (\Box P \rightarrow \Box Q)$ (K-schema)

A3. $\Box P \rightarrow \Diamond P$ (D-schema)

R1. $\vdash P, \vdash P \rightarrow Q$ then $\vdash Q$ (Modus Ponens)

R2. $\vdash P$ then $\vdash \Box P$ (necessity closure)

The D-schema formalizes the principle that “whatever is obligatory is permissible”.

Semantically, it is the requirement that the relation R be serial that underwrites the D-schema.

Proposition 1.0 Let $F = \langle W, R \rangle$ be a frame. Then R is serial if, and only if, $\Box P \rightarrow \Diamond P$ is valid in F .

For the proof of prop 1.0, I will introduce the following definition of a propositional modal *Model*:

Definition 1.5 A modal propositional *model* is an ordered triple $M = \langle W, R, I \rangle$, where

1. $\langle W, R \rangle$ is a frame
2. I is a function which assigns, to any pair consisting of a world from W and a statement letter A , exactly one truth value, either True or False.

We can thus characterize the valid formulas to be those formulas which are true at all worlds under all interpretations. Given a frame, $F = \langle W, R \rangle$, $M = \langle W, R, I \rangle$ is called a model based on F . Therefore, when supplied with a model based on F , we then have a structure which provides truth values for all atomic statements at each world in F . We extend the interpretation to all formulas as follows:

Definition 1.6 P is true at w in model M , i.e., for world w , $M \models_w P$ when,

1. If P is atomic, $M \models_w P$ iff $I \langle w, P \rangle = \text{true}$, i.e., P is a member of w .
2. $M \models_w P \rightarrow Q$ iff, if $M \models_w P$ then $M \models_w Q$, i.e., $M \not\models_w P$ or $M \models_w Q$
3. $M \models_w \sim P$ iff, $M \not\models_w P$
4. $M \models_w \Box P$ iff for any $x \in W$, if wRx then $M \models_x P$
5. $M \models_w \Diamond P$ iff there is a $x \in W$ such that wRx and $M \models_x P$

Proof of Prop 1.0: Given a serial F , suppose that $M \models_w \Box P$. Thus, there is an x such that wRx and $M \models_x P$. Consequently, $M \models_w \Diamond P$, as desired. From right to left, assume that $\Box P \rightarrow \Diamond P$ is a valid formula in F . However, assume that there is no x such that wRx ; then trivially P is true at every x accessible from w , so $M \models_w \Box P$ but not $M \models_w \Diamond P$, contradicting our supposition. **Q.E.D.**

We define \Box° , \Diamond° , \Box^2 and \Diamond^2 thusly:

Definition 1.7 ' $M \models_w \Box^\circ A$ ' is true iff it is true that for any x , if $wR^\circ x$ then $M \models_x A$.

Definition 1.8 ' $M \models_w \Diamond^\circ A$ ' is true iff it is true that there is at least one x , such that $wR^\circ x$ and $M \models_x A$.

Definition 1.9 ' $M \models_w \Box^2 A$ ' is true iff it is true that for any x , if $wR^2 x$ then $M \models_x A$.

Definition 1.10 ' $M \models_w \Diamond^2 A$ ' is true iff it is true that there is at least one x , $wR^2 x$ and $M \models_x A$.

As noted earlier, it is the *seriality* of the accessibility relation R° which is seen as the required property of a good formalization of equating moral necessity with the quantification over morally relevant localities.

However, in order to validate the inference from ‘obligation’ to ‘can’ with a polymodal logic combining **KD** and **S5**, the two R -relations must be related to each other as follows:

Proposition 1.1 The class of bi-modal frames $\langle W, R^\circ, R^2 \rangle$ must be such as to provide a non-empty W , R° being a serial relation on W , R^2 being an equivalence (i.e., reflexive, transitive, and symmetric) relation on W , and for any y , if $wR^\circ y$ then $wR^2 y$. This condition results in the following formulas being valid in such a frame: $\Box^\circ P \rightarrow \Diamond^2 P$; $\Diamond^\circ P \rightarrow \Diamond^2 P$.

The former formula may be seen to represent the claim that ‘an obligation can always be discharged’, while the latter formula may be seen to represent the claim ‘that which is permitted is always possible’.

Proof of Prop 1.1: Suppose that at a world x , $\Box^\circ A$ holds. As noted above, for any w , if $xR^\circ w$ then $xR^2 w$; and since for any w , such that $xR^\circ w$, then $w \Vdash A$. Thus, for at least one w , such that $xR^2 w$, $w \Vdash A$. Therefore $x \Vdash \Diamond^2 A$, as desired. Conversely, suppose that $\Box^\circ A \rightarrow \Diamond^2 A$ is valid in the frame and that w is a world in the frame for which there is no x , such that $wR^2 x$ but that x is nonetheless $wR^\circ x$ and for any x , $x \Vdash A$. Thus we have $w \Vdash \Box^\circ A$ but not $w \Vdash \Diamond^2 A$ contradicting our supposition. For $\Diamond^\circ A \rightarrow \Diamond^2 A$: Suppose that $M \Vdash_x \Diamond^\circ A$. By definition of \Diamond° , there is a y , $xR^\circ y$ and $M \Vdash_y A$. According to the rules governing R° , R^2 interaction, for any y , if $wR^\circ y$ then $wR^2 y$; thus this y is $xR^2 y$. Since this y $M \Vdash_y A$, then we have $M \Vdash_x \Diamond^2 A$ as desired. Conversely, suppose that $\Diamond^\circ A \rightarrow \Diamond^2 A$ is valid in the frame and that w is a world in the frame for which there is no x , such that $wR^2 x$ but that x is

nonetheless $wR^{\circ}x$ and $x \vDash A$. Thus we have $w \vDash \diamond^{\circ}A$ but not $w \vDash \diamond^2A$, contrary to our supposition.

SDL and Genuine Moral Dilemmas

As an example of an advocate for equating moral obligation with moral necessity (classically understood), Terrance McConnell proposes that the conjunction of the following three sentences express an inconsistent set of propositions: (1) there are genuine moral dilemmas; (2) *ought* implies *can*; (3) if one is obligated to do each of two courses of action, then one is obligated to do both courses of action. McConnell uses the following argument, in his own multi-modal take on standard deontic logic (SDL), to show that the above set of sentences is inconsistent. In his version of SDL, **OA** stands for “it is obligatory that A” and \diamond stands for classical alethic possibility.

- | | |
|---|---|
| 1) OA | premise |
| 2) OB | premise |
| 3) $\sim \diamond (A \ \& \ B)$ | premise (1-3 represent a moral dilemma) |
| 4) $\mathbf{O}(A \ \& \ B) \rightarrow \diamond (A \ \& \ B)$ | premise; by way of (2) |
| 5) $\mathbf{(OA \ \& \ OB)} \rightarrow \mathbf{O(A \ \& \ B)}$ | premise; by way of (3) |
| 6) O(A&B) | 1,2,5 &I, \rightarrow E |
| 7) $\sim \mathbf{O(A \ \& \ B)}$ | 3,4 m.t. |

McConnell goes on to argue that the fault for this inconsistency lies with the assumption that there can be genuine moral dilemmas. His argument mostly rests on the notion that the formulas in #4 and #5 in the argument above are axioms of standard deontic logic. Thus he suggests that the following formula in deontic/alethic polymodal logic may be

interpreted to represent a “condition of adequacy for any ethical theory [to] not allow for genuine moral dilemmas”:

$$(\mathbf{OA} \ \& \ \mathbf{OB}) \rightarrow \diamond(A \ \& \ B)$$

I understand McConnell as arguing that $\sim [(\Box^\circ A \ \& \ \Box^\circ B) \ \& \ \sim \diamond^2(A \ \& \ B)]$ is a valid formula of a formal system, which when appropriately interpreted using deontic and alethic terms, expresses the proposition that there are no genuine moral dilemmas.

However, recall from earlier that we needed to introduce a semantic postulate in order to characterize the appropriate class of bi-modal frames $\langle W, R^\circ, R^2 \rangle$. Such frames must be such as to provide a non-empty W , R° being a serial relation on W , R^2 being an equivalence (i.e., reflexive, transitive, and symmetric) relation on W , and for any y , if $wR^\circ y$ then $wR^2 y$. We previously saw that this condition results in the following classically desirable formulas being valid in such a class of frames: $\Box^\circ P \rightarrow \diamond^2 P$; $\diamond^\circ P \rightarrow \diamond^2 P$. While these formulas are desirable for classical theorists, the above semantic postulate also makes valid: $\Box^2 P \rightarrow \Box^\circ P$, which may be seen to represent the somewhat intuitively odd claim, that if something is necessarily true, then it is morally obligatory for it to be true. Moreover, this claim may, arguably, be seen as a version of the naturalistic fallacy: deriving an ‘ought’ from an ‘is’, albeit a ‘necessary is’.

Proof: Suppose that $x \not\models \Box^2 P$. Thus, we have that for any y such that $xR^2 y$, then $y \not\models P$. Suppose also that $x \not\models \sim \Box^\circ P$. Therefore there is a y , such that $xR^\circ y$ and $y \models \sim P$. And with our semantic postulate, this leads to contradiction.

Such a result may indicate to some theorists that SDL, in the formalization provided by the polymodal **KD** and **S5**, is too strong to be a good formalization of deontic

reasoning. But for all we have shown so far, perhaps the impossibility of a formalization of genuine moral dilemmas resides in this interaction of R° and R^2 .

We can see that this is not true by considering **KD**. In **KD**,

- 1) $\Box(A \& B) \rightarrow \Diamond(A \& B)$
- 2) $(\Box A \& \Box B) \rightarrow \Box(A \& B)$
- 3) $\Diamond(A \& B) \rightarrow (\Diamond A \& \Diamond B)$
- 4) $\sim(\Box A \& \Box \sim A)$

Proof 1): Assume $x \Vdash \sim[\Box(A \& B) \rightarrow \Diamond(A \& B)]$. Thus, $x \Vdash \Box(A \& B) \& \sim\Diamond(A \& B)$. We have $x \Vdash \sim\Diamond(A \& B)$ by $\&e$, and thus infer that $x \Vdash \Box \sim(A \& B)$ by \Box - \Diamond interchange. The D-schema gets us $x \Vdash \Diamond \sim(A \& B)$. Thus there is a y such that xRy and $y \Vdash \sim(A \& B)$. Since $x \Vdash \Box(A \& B)$, we also infer that $y \Vdash (A \& B)$ holds as well, which results in contradiction.

Proof 2): Assume $x \Vdash \sim[(\Box A \& \Box B) \rightarrow \Box(A \& B)]$. So: $x \Vdash (\Box A \& \Box B) \& \sim\Box(A \& B)$. We have $x \Vdash \Box A$, $x \Vdash \Box B$ from two applications of $\&E$. From interchange on $x \Vdash \sim\Box(A \& B)$, we infer $x \Vdash \Diamond \sim(A \& B)$. Thus there is a y such that xRy and $y \Vdash \sim(A \& B)$. From $x \Vdash \Box A$, $x \Vdash \Box B$ infer $y \Vdash A$ and that $y \Vdash B$ and therefore that $y \Vdash A \& B$, which leads to contradiction.

Proof 3): Assume that $x \Vdash \sim[\Diamond(A \& B) \rightarrow (\Diamond A \& \Diamond B)]$. So: $x \Vdash \Diamond(A \& B) \& \sim(\Diamond A \& \Diamond B)$. By $\&E$, $x \Vdash \sim(\Diamond A \& \Diamond B)$. Thus we have either $x \Vdash \sim\Diamond A$ or we have $x \Vdash \sim\Diamond B$. Thus we may suppose either $x \Vdash \Box \sim A$ holds or that $x \Vdash \Box \sim B$ holds, by interchange of $\sim\Diamond$. We have $x \Vdash \Diamond(A \& B)$ by $\&E$, as above. Therefore there is a y such that xRy and $y \Vdash A \& B$. But from either case, $x \Vdash \Box \sim A$ or $x \Vdash \Box \sim B$, we now end up in contradiction at y .

Proof 4): Suppose that $x \Vdash \Box A \& \Box \sim A$. Then we have $x \Vdash \Box A$, by $\&E$. The D-schema nets us $x \Vdash \Diamond A$. Therefore there is a y such that xRy and $y \Vdash A$. Since we also have $x \Vdash \Box \sim A$ by $\&E$ we infer that at any y such that xRy then $y \Vdash \sim A$, which leads to contradiction.

And of course, given the classical definition of negation, we can't both have an obligation to do A and not have an obligation to do A in **KD**.

Therefore we have:

Proposition 1.2 Genuine moral dilemmas are still impossible according to **KD**—that is, if A and B are logically incompatible then $(\Box A \ \& \ \Box B)$ is unsatisfiable:

Proof of Prop 1.2: $\sim(\Box P \ \& \ \Box \sim P)$ is valid in **KD**, as shown in proof 4 above.

$\sim(\Box P \ \& \ \sim \Box P)$ is valid in **KD**, by definition of negation.

If 'A; B $\not\models f$ ', where f represents any contradiction, is true, then $\sim(\Box A \ \& \ \Box B)$ is verified by any model in **KD**: Let us suppose that 'A; B $\not\models f$ ' is true in a model. Let us also suppose that $x \not\models \Box A$ and $x \not\models \Box B$ and therefore that $\forall y$, if xRy then $y \not\models A$, and $y \not\models B$. By seriality, there is a w such that xRw . So: there is a w such that $w \not\models A \ \& \ B$, which according to our assumption, is impossible.

Standard Deontic Logic (SDL) and its affiliates cannot represent genuine moral dilemmas. In fact, some proponents of SDL explicitly argue that SDL is incompatible with genuine moral dilemmas. As we have seen, they are right about this. However, how is it possible for proponents of such logics to reasonably reconcile this fact with the fact that moral agents all too frequently find themselves, at least on first sight, in positions of conflicting obligations? I propose it is impossible for proponents of SDL to reasonably reconcile its no-dilemmas feature with moral experience, unless they cherry-pick what counts as genuine elements of moral experience.

The advocate of SDL may object to the charge of cherry-picking, by claiming that, of whatever sets of obligations populate the ethical domain, it is true that it is

impossible for them to be inconsistent. However, let us press on, and see whether this claim is at all plausible after considering only one type of obligation: promise-making.

Consider how few theorists will argue that someone, either through his or her own wrongdoing or ignorance, could not make incompatible promises. For example, consider Jack who, knowing that he can't be in two places at the same time, nonetheless promises Felicity that he will meet her in London at 3pm and promises Vera that he will meet her in Toronto at 3pm that same day. Therefore, if keeping promises are obligatory, as is commonly assumed, there is at least one class of self-inconsistent sets of obligations, contrary to the initial claim of SDL proponents.

Nevertheless, proponents of SDL and most informal theorists may see fit to revise the above claim. They may argue that SDL proves that it is only obligations incurred without your own wrong-doing or ignorance so corrupting the situation, which are always consistent. This is exactly the type of cherry-picking that SDL requires to defend it. SDL requires postulating that some sets of moral obligations are somehow outside the domain of ethical discourse. Unless there is independent support for such a postulate, SDL deserves to be seen as cherry-picking the characteristics of appropriate moral inquiry. Until such independent support can be procured, it seems reasonable to require that a good formalization of deontic logic should not preclude an inconsistent set of obligations from being a part of an appropriate moral inquiry.

To add insult to injury, consider that it is not always your own wrongdoing or ignorance which may be responsible for you making incompatible promises or incurring incompatible obligations. Let us thus suppose that Ronald, who happens to be a very bad person, tells Jack that St. Mary's Church is located on the corner of 12th street and Jester

Avenue, which it isn't. Jack then, in good faith, promises Felicity to meet her at St. Mary's Church at 3pm and promises Vera to meet her on the corner of 12th street and Jester Avenue at 3pm, believing that he can discharge both obligations as incurred. Now, it doesn't seem correct to insist that Jack is to blame for the incompatible character of his obligations, but rather it is the evil Ronald who is to blame, since he lied to Jack about the location of the church.

If we then argue that Ronald's behavior doesn't likewise cause any trouble for those informal theories wedded to the claim that all of Jack's 'normal' sets of obligations may be realized, it seems that we may be merely begging the question: we recognize an acceptable (i.e., 'normal') set of obligations only if the set may be coherently or jointly realized. Again we are merely cherry-picking what counts as genuine moral experience in order to make the SDL characterization of consequence fit moral inquiry.

Some may object that it is still Jack's own ignorance of the true location of the church which is truly responsible for the incompatibility of his obligations; or perhaps Jack did wrong in trusting Ronald's regarding the whereabouts of the church. Even if these objections get it right, consider how many real obligations could be incurred under the ideal conditions of always knowing who to trust and always knowing whether responsibilities we choose to undertake will be jointly realizable.

To make the implicit principle to which I am appealing explicit, consider the following analogy to the *principle of epistemic humility*: in similar fashion to how humans are epistemically fallible, and thus no one can reasonably think that every belief he or she has is true, humans are morally fallible, and inevitably so. Therefore, no matter how careful we are, we can never be sure that we or others like us haven't done or will do

something morally wrong or that we are not ignorant of some morally relevant feature, which may in turn, render our own or others' future set of obligations unsatisfiable.

In virtue of this similarity to epistemic humility, I will call this principle the *Principle of Deontic Humility*. Together with epistemic humility, the truth of deontic humility renders the inference from 'ought' to 'can' untenable: we can never know that there is no prior wrongdoing, committed by oneself or by someone else, which precludes the satisfaction of any given set of obligations. Indeed, epistemic humility means that a reasonable person should think that any one of his or her beliefs may be false; deontic humility guarantees that there is 'stuff' we can be wrong about, namely, whether or not there is some prior wrongdoing which precludes the satisfaction of a given set of obligations. This consideration seems to imply that deontic humility is entailed by epistemic humility; however, I will not pursue this issue here.

The above argument may be outlined thusly, with X representing any arbitrary action as long as it is possible for some wrongdoing to preclude someone from doing it⁴:

- 1) If you ought to do X then necessarily, you can do X.
- 2) If you can do X, then necessarily, there are no prior wrongdoings which precludes doing X.
- 3) Humans inevitably fail to always do what is morally right. If there could be some prior wrongdoing which could preclude doing X, it is unreasonable to suppose that it is impossible that no one has done that wrongdoing.
- 4) Therefore it is unreasonable to suppose that it is always possible that you can do X...thus it is not the case that for any X, it is reasonable to suppose that you ought to do X.

Thus, unless no action which may be precluded by some wrongdoing is ever obligatory, we must reject either that 'ought' implies 'can' or the *Principle of Deontic Humility*.

Since the *Principle of Deontic Humility* is grounded in ordinary moral experience and

⁴ Otherwise, the claim that 'ought' implies 'can' would seem rather mundane.

common-sense, rejecting the *Principle of Deontic Humility* comes at a far higher cost than does rejecting idealized conceptions of morality which always permit the inference from ‘ought’ to ‘can’.

Therefore, I propose that the class of such classes of obligations which, in virtue of certain conceptual truths regarding obligation and its cognates, are necessarily compatible and jointly realizable will likely be quite small, perhaps even empty⁵. In other words, such conceptions may well be inapplicable to our world of experience: they fail to accurately describe the features that are part of the actual world because these theories contravene the *Principle of Deontic Humility*.

Secondly, I propose that genuine moral dilemmas deserve to be treated, *prima facie*, as elements in moral agents’ lives. The experiences of impossible moral situations, in which the values justifying incompatible prescriptions are incommensurable, are a mainstay in literature and cultural products, from the sublime in Sartre to the low-brow and juvenile in comic-book heroes. I propose that the apparent success of such cultural products is due to the resonance most consumers experience when they consider the moral quandary in which the protagonists find themselves confined⁶. Moreover, it is unfortunately far too easy to imagine that cases even worse than Searle’s choice occur too often in our world.

My argument is conductive in nature; however I believe it permits one to reasonably believe, contrary to the claims of SDL’s proponents, that moral agents can

⁵ Consider the implications that the belief in a devil might have on this line of reasoning: that it is likely that all or nearly all our obligations we might incur will have historical roots in the activities of an evil agent much worse than our evil Ronald. For those who argue for the no-dilemmas position for theological reasons, it might well behoove them to reconsider the appropriateness of such a position if their ontology includes such evil forces.

⁶ Furthermore, in the case of comic book heroes, it is only in virtue of their super-powers that they are able to reconcile what would be for the rest of us, incompatible obligations.

and do experience and reason from genuine moral dilemmas. The belief that all sets of obligations are compatible and jointly realizable is, I propose, a matter of faith in the applicability of certain conceptual truths to our world of experience. Contrarily, the complexity and diversity of our moral lives, together with the nearly common, shared experience of moral dilemmas, provides a *prima facie* case against the appropriateness of conceiving obligation in the manner proposed by advocates of SDL and its affiliate informal ethical theories.

KD and its Affiliates Tackle Trumping

With the D-schema, we infer from $x \models \Box A$ that $x \models \sim\Box\sim A$. However, when we say that another obligation, let's say $\Box B$, trumps $\Box A$, we say that it is now impermissible to do A if doing A precludes also doing B. Therefore, adding new premises changes what could be originally inferred from $x \models \Box A$, namely that $x \models \Diamond A$. Thus, at least at first blush, monotonic logics cannot provide an adequate formalization of moral theories which contain principles permitting one obligation to trump another obligation.

SDL and **KD** are types of monotonic logic. For at each locality (i.e., accessible worlds), truth-functional formulas are given truth-values using classical truth tables⁷. And the truth-functional formulas of classical logic are widely known to be monotonic. Furthermore, recall that in **KD**, we know that for any x in any interpretation if $x \models \Box A$ then $x \models \Diamond A$. Let us suppose that $x \models \Box A$. Thus we infer $x \models \Diamond A$. We show that in **KD** for any x , if $x \models \Box A$ and $x \models \Box B$ then $x \models \Diamond A$ also. Suppose now that not only $x \models \Box A$ but that $x \models \Box B$ also. Nonetheless, for any x in any interpretation if $x \models \Box A$ then $x \models \Diamond A$, thus we have $x \models \Diamond A$ regardless of the addition of any further premises.

⁷ This is a sufficient condition for monotonicity, given that the R accessibility relation is the same for any arbitrary x and w such that wRx .

Consequently, the inference that if $x \models P$ then $x \models Q$ therefore if $x \models P \ \& \ R$ then $x \models Q$ is thus valid for any x in any interpretation according to **KD** and its affiliates. Thus, adding new premises cannot change what was previously implied by the original set of premises in **KD**. This feature of **KD**, I argue, renders it inappropriate to be applied to moral reasoning. By ‘trumping’ we mean specifically a type of reasoning, which by introducing new premises, changes what was previously implied by the original set of premises.

Not only does **KD** fail to be non-monotonic, I will show in what follows that **KD**, when interpreted using moral terms, can not express the notion of a hierarchy of sets of obligations taking part in an appropriate moral inquiry. Informally, trumping-like reasoning only makes sense when we must know which member(s) of an inconsistent set of obligations over-rides the others. **KD** explosively infers that anything and nothing is obligatory from an inconsistent set of obligations: Obligation Explosion. If this holds for **KD**, then it likewise must hold for a multimodal **KD** and **S5**: **SDL** and its affiliates. Therefore **SDL** cannot represent trumping-like reasoning.

We now turn to formally investigating whether **KD** is capable of representing how one obligation can over-ride another obligation. To do so, we introduce a new operator \prec , a binary relation on sets of obligations: we shall write $x \models P \prec Q$, for the situation such that whenever an obligation from type Q comes into conflict with an obligation from type P , the obligation from type Q should have more weight in moral reasoning than obligations from P , or alternatively, that Q -type obligations trump P -type obligations. Hierarchies then can be understood as containing a possibly partial ordering of sets of obligations given by \prec .

Proposition 1.3 *KD* and any of its affiliates cannot, without explosively proving that everything and nothing is obligatory, represent moral reasoning from $P; Q \not\models f$ together with a hierarchy, which returns with a desirable result: for example provided that $x \not\models Q < P$, if $x \models \Box Q$ and $x \models \Box P$ then $x \models \Box \sim Q \ \& \ \Box P \ \& \ \Diamond P \ \& \ \Diamond \sim Q$.

To prove this, we define a hierarchy:

Definition 1.11 A hierarchy h is here stipulated as a function on sets of obligations. Each set is self-consistent, i.e. contains only jointly realizable obligations. At point x , h : for any A, B such that $[((\Box A \ \& \ \Box B) \text{ but } A; B \not\models f) \ \& \ \text{for some } P, Q \ ((\Box A \in P \ \& \ \Box B \in Q) \vee (\Box A \in Q \ \& \ \Box B \in P)) \ \& \ \sim((\Box A \in P \ \& \ \Box B \in P) \vee (\Box A \in Q \ \& \ \Box B \in Q))] \ \& \ (P < Q \vee Q < P \ \& \ \sim(P < Q \ \& \ Q < P)^8)] \rightarrow [((Q < P \ \& \ \Box A \in P) \rightarrow \Box \sim B) \vee ((Q < P \ \& \ \Box B \in P) \rightarrow \Box \sim A) \vee ((P < Q \ \& \ \Box A \in Q) \rightarrow \Box \sim B) \vee ((P < Q \ \& \ \Box B \in Q) \rightarrow \Box \sim A)]$.

Proof of Prop 1.3: Assume $x \models \Box A$. Introduce new premise $x \models \Box B$, but that $A; B \not\models f$.

Suppose that $\Box B$ belongs to P & $\Box A$ belongs to Q and a hierarchy assigns $Q < P$. Then we infer $x \models \Box \sim A$ using the hierarchy function. Since *KD* $\vdash \sim(\Box A \ \& \ \Box \sim A)$, if $x \models \Box A$, $x \models \Box \sim A \vee \Box C$, for some arbitrary C , then by disjunctive syllogism, we may infer that anything and nothing is obligatory or permissible.

Even without the above characterization of a hierarchy, any informal trumping-like reasoning along the lines of:

‘suppose that $x \models \Box P$ and $x \models \Box Q$ but that $P; Q \not\models f$, so we need action-guiding. Our favorite informal theory claims that $\Box Q$ trumps $\Box P$, therefore we must do $\Box Q$ at the expense of $\Box P$ ’,

...results in obligation explosion. So:

⁸ Since $<$ is symmetrical and we are considering partial orders, then if $P < Q$ and $Q < P$ then $P = Q$. This is undesirable given that we want members of one set of obligations to trump other, lower ranked obligations.

Proposition 1.4 *KD* is incompatible with any trumping-like reasoning, if we do not desire to verify the truth of every formula.

Proof: This can be seen straightforwardly: Suppose that $P; Q \not\models f$, but that $x \models \Box P$ and $x \models \Box Q$, in order to set up a situation which calls for trumping-like reasoning. However, we infer that for any y such that xRy , $y \models P$ and $y \models Q$. By our assumption, $P; Q \not\models f$, and since we can infer $y \models P \vee M$ from $y \models P$, we infer that $y \models M$, for any arbitrary M . By \Box -definition we have $x \models \Box M$, for any arbitrary M .

There is nothing to block this result within *KD* unless one rejects the possibility that an obligation can be overridden by a later obligation. This seems to undercut the basis for the ‘action-guiding’ criterion of good informal theories. Thus the proponent of Standard Deontic Logic is incapable of using any logic sufficiently similar to *KD* to represent the sentences of his or her favorite informal ethical theories, if such theories are considered ‘action-guiding’ in this sense.

Again, the valid formula of *KD*, $\sim [(\Box A \ \& \ \Box B) \ \& \ \sim \Diamond(A \ \& \ B)]$, when interpreted with deontic terms, with or without *S5*, rules out any and all incompatibility of moral obligations. Moreover, any hierarchical ranking of sets of obligations in which one obligation trumps another obligation, seen as so essential to action-guiding in informal ethical theories, is also excluded by this formula⁹. Thus, when applied to informal moral reasoning, *KD* and its affiliates should not be acceptable as a good formalization.

With *KD*, in which moral obligations cannot conflict, if two demands do conflict, one of them could never have been an obligation in the first place, regardless of its

⁹ That is, given the monotonicity of *KD* and the undesirable consequence of everything and nothing being obligatory

supposed moral virtues. In other words, unless the possibility that everything and nothing is obligatory is countenanced, nothing should be considered an obligation until it can be known to cause no conflict with any other higher ranked obligation we may incur. Such an epistemic requirement seems far too high of a cost in return for the simplicity of a classical formalization of deontic reasoning.

On the other hand, perhaps all we are left with is *prima facie* obligations. But to accept that obligations generally have only *prima facie* status seems to come at a steep cost: what sense can be made of a *prima facie* moral necessity, classically understood? The notion of a *prima facie* moral necessity is, in classical modal logic, incoherent; thus treating obligations as having *prima facie* status only does harm to the plausibility of the SDL proponent's cause.

Making Valid the Naturalistic Fallacy

Consider again how SDL permits the inference:

$$\frac{\vdash A}{\vdash \Box A}$$

However, as Mares (2004) shows, this closure rule, together with permitting the addition of irrelevant premises and the deduction theorem (or \rightarrow I) we can prove:

$$\frac{\frac{\frac{\vdash A \text{ (hypothesis)}}{\vdash \Box A \text{ (closure under } \Box)}{p \vdash \Box A \text{ (weakening)}}}{\vdash p \rightarrow \Box A \text{ } (\rightarrow\text{I})}$$

It seems reasonable to read $\vdash p \rightarrow \Box A$ as permitting, as a matter of logic, the inference of an 'ought' from an 'is', when applied to deontic logic. I imagine that many philosophers will find it undesirable that SDL so straightforwardly makes the naturalistic fallacy valid.

I propose that this result provides at least an ancillary justification, together with SDL's problems with representing genuine moral dilemmas and trumping, for rejecting SDL as a good formalization of ethical discourse. Therefore, the following chapters will focus on the comparative merits using two formal, non-monotonic and para-consistent systems, a relevant modal logic, and a preservationist-style modal logic, in order to express the sentences of any adequate informal ethical theory.

Chapter Two

A Relevant Deontic Logic: Considering the Impossible

When we previously considered the fit between SDL and the set of inferences we desire to be valid in ethical reasoning, SDL's failures appear to be in part due to its inclusion of the fallacies of irrelevance in its set of valid formulas. SDL permits explosion, which entails that an impossible moral situation cannot be used fruitfully in moral reasoning (since a contradiction implies anything). Likewise, the addition of irrelevant premises is permitted in SDL's characterization of the inference relation. The validity of this fallacy of irrelevance in SDL causes two seemingly insurmountable problems. Firstly, we saw that the monotonicity of this inference relation torpedoes any attempt to formalize one obligation trumping another. Secondly, SDL, due to weakening and necessitation closure makes the naturalistic fallacy valid.

Situations and Worlds

In the relevant logic **R** (for relevant implication), the class of models capable of being used to characterize the class of validities will have to be more general than its classical counterpart: we need to increase the set of objects to quantify over in order to enlarge the class of counter-examples. Recall that, as shown in the previous chapter, for a classical frame: $F = \langle W, R \rangle$, W is a non-empty set of possible worlds; furthermore, if each point in a model: $M = \langle W, R, I \rangle$ is characterized as a possible world, then according to any classical semantics, such possible worlds deserve to be understood as "deciding every issue"¹⁰. Thus, for any proposition, each point belongs either to the set of points for which that proposition receives the value *true* or to the set of points for which that proposition receives the value *false*. Likewise possible worlds are completely consistent:

¹⁰ Pg 27 Mare

for any possible world x , it is never the case that $x \models p \ \& \ \sim p$. The indices in frame-semantics for a relevant logic, alternatively, should be understood as being more like possible ‘situations’ rather than possible worlds.

Worlds decide every issue; situations, on the other hand, may lack information regarding whether an arbitrary proposition is true or false. Furthermore, no possible world in a model can belong to a set which makes both \mathbf{P} and $\sim\mathbf{P}$ true, i.e., possible worlds are completely consistent; some situations, however, may make such contradictions true¹¹. We will call situations which lack information, ‘partial situations’, and situations that are inconsistent, ‘inconsistent situations’; those situations that are complete and consistent, we will call ‘logical situations’ and those situations which are either partial or inconsistent (or both) are called ‘non-logical situations’.

Logical situations have a tidy relationship with validity in a relevant logic model. Those formulae verified by the empty set are true at all logical worlds; this is exactly what we should expect: logical truths will be true at logical situations. However, there is no guarantee, and reasonably so, that the valid formulae of any logic will be true at the non-logical situations in its semantics.

The first section in this chapter will explicate and justify using such situations in modeling the class of valid deontic inferences. But before so doing, I will briefly introduce a ternary accessibility relation R on situations, which is integral for the following characterization of the relevant truth condition for implication. Together with definitions of a persistence relation, relevant frames and models, and a semantic version of a deduction theorem, we will be able to set up the logical machinery for the positive fragment of relevant implication: $\mathbf{R}+$.

¹¹ *ibid*

Definition 2.0 (Relevant implication) $x \models A \rightarrow B$ in a model iff for any y and for any z in the model if $(Rxyz \ \& \ y \models A)$ then $z \models B$

To explicate this definition, we first need to stipulate the properties of a *persistence* relation on situations, offer a semantic version of the deduction theorem, and provide definitions of a relevant frame and model

Definition 2.1 A persistence relation is a binary relation on situations, such that for any situations t and u , u extends t iff there is at least one logical situation s , and $Rstu$. When u extends t , we say that t is a part of u ; thus in order theory terminology, $t \leq u$. In words, suppose we have a set of information due to a logical s , and that $Rstu$. If we were to hypothesize that something from situation t obtains in the same world as s , then we would be entitled to infer that something from situation u would also obtain in the same world. Thus, if $Rstu$, and s is logical, then $I(t) \subseteq I(u)$

In order to characterize an appropriate deduction theorem, we will need persistence to be a partial order: reflexive, transitive, and anti-symmetrical. Furthermore, if an interpretation I assigns *true* to any proposition p at t and u extends t , then I must also assign *true* to p at u , in virtue of the part-whole relationship between these situations.

Definition 2.2 A relevant frame is an F such that $F = \langle sit, logical, R \rangle$, where sit is a non-empty set of situations, $logical$ is a non-empty subset of sit , the set of logical situations¹², and an accessibility relation R , which possesses the following properties:

For any s, t, u of sit

1. If $Rstu$, then $Rtsu$ (interchange)

¹² That is situations which obey all the classical laws of logic...analogous to possible worlds

2. If $Rstu$, then $Rsst$ (repetition)
3. $Rsss$ (complete reflexivity)
4. If $Rstu$ and s extends s' , then $Rs'tu$ (transitivity)

Definition 2.3 A relevant model is an M such that $M = \langle sit, logical, R, I \rangle$ where $\langle sit, logical, R \rangle$ is a relevant frame and I is an interpretation which takes situations into the power-set of atoms, and thus assigns a situation s to $I(p)$ if p is true at s according to I .

Theorem 2.0 Semantic Entailment: If the relevant entailment from 'A' to 'B' is verified in a model then 'A \rightarrow B' is likewise verified at all logical situations in the model.

Moreover, if an entailment is verified in any model, then its subsequent implication is likewise verified at all logical situations in all models. Therefore such an implication would necessarily be a valid formula of the system.

Proof: Given a model M , suppose we have formulae 'A' and 'B', an arbitrary logical situation s and arbitrary situations t and u , such that $Rstu$. Assume that for any arbitrary situation y , if M verifies formula 'A' at y , then M also verifies 'B' at y . Since s is a logical situation and $Rstu$, then u extends t . Let us suppose that M verifies 'A' at t ; according to our assumption, M also verifies 'B' at t . Thus, M likewise verifies 'B' at u due to the persistence relation. Consequently, we have $s \models A \rightarrow B$, since for any y and for any z in the model if $(Rsyz \ \& \ y \models A)$ then $z \models B$. Furthermore, according to our assumption, we may, for any situation, deduce B from A. Thus, at s we may deduce B from A. **Q.E.D.**

The Positive fragment¹³ of logic R

With the above definitions of a relevant frame, model and a ternary accessibility relation, the truth conditions for formulae without negation (for \rightarrow , \vee , $\&$, for any

¹³ The logic of Relevant implication without negation

propositional variable p and formulae A and B , where ' $s \models A$ ' means that A is true at s according to I) may be given as follows:

1. $s \models p$ iff $s \in I(p)$
2. $s \models A \ \& \ B$ iff $s \models A$ and $s \models B$
3. $s \models A \vee B$ iff $s \models A$ or $s \models B$ ¹⁴
4. $s \models A \rightarrow B$ iff for any x , for any y , if $(R_sxy \ \& \ x \models A)$ then $y \models B$

A Relevant Negation

The truth conditions for negation in classical logic, given a standard semantics, may be characterized as being analogous to set compliment:

Definition 2.4 Classical negation for propositional variable p , world x and interpretation I : ' $x \models \sim p$ ' iff x does not belong to the set of possible worlds in which p is true according to I .

Classical negation for formula ' A ', world x and interpretation I : ' $x \models \sim A$ ' iff ' A ' fails to be true in x .

By equating negation with the failure of points in a model to belong to the set of points in which the proposition or formula is true seems to entail, by all appearances¹⁵, that each point is complete and consistent. With models in which all points are complete and consistent, many of the fallacies of irrelevance may be verified, as demonstrated in many

¹⁴ Note that an alternative clause may be desirable for disjunction: $s \models A \vee B$ iff there is an x such that Msx and $x \models A$ or $x \models B$. M is in this case an binary accessibility relation on sit ; this may be read as claiming that there is an accessible, but (possibly) as yet not accessed, situation in which either A will be found to be true or B will. For example, consider that, given what we know now (call this situation s) about a well-confirmed scientific theory, together with knowledge that with one more piece of information that is soon forthcoming (call this situation t), it is only a matter of time before we discover whether A is the case or B is the case. Thus at s , even before we receive that one more piece of information, we know that either A is the case or B is the case. So it is not the case that $s \models A$ or $s \models B$, but $s \models A \vee B$ nonetheless.

¹⁵ On the other hand, Graham Priest claims that this inference depends on the classical metalanguage lurking in the background, and once rejected, the 'fails to be true' clause does not entail complete and consistent points.

accessible primers in relevant logic. Thus, the classical semantics for negation need to be jettisoned for the current project.

To define a relevant negation, consider that, unlike classical semantics, we have a class of impossible points. This class may be equated with the proposition expressed by “something impossible occurs”¹⁶ which we will call proposition f ¹⁷; so that $I(f) =$ impossible situations, i.e., $\{x \mid x \not\models I P \ \& \ \sim P\}$.

Definition 2.5 Relevant negation:

‘ $s \models I \sim A$ ’ is true iff for any x and for any y , $(R_s x y \ \& \ x \models I A) \rightarrow y \models I f$.

This definition, with a model $\mathbf{M} = \langle sit, logical, f, R, I \rangle$, where f as defined above, is a subset of sit , gives us negation in terms of a relevant implication of an impossible situation being realized: $\sim A = A \rightarrow f$. Furthermore, we can now define both an incompatibility and compatibility relation between situations.

Definition 2.6 With a model $\mathbf{M} = \langle sit, logical, R, N, I \rangle$ an incompatibility relation N is a binary relation on sit such that Nst iff for any x if $Rstx$ then $x \not\models f$

Note that a situation may be incompatible with itself. For example if $t \models I A \ \& \ \sim A$, then Ntt .

Definition 2.7 With a model $\mathbf{M} = \langle sit, logical, R, C, I \rangle$ a compatibility relation C is a binary relation on sit such that Cst iff for any x if $Rstx$ then it is not the case that $x \not\models f$

We now have two more ways of formalizing negation, in terms of a compatibility relation or an incompatibility relation:

Definition 2.8 Negation in terms of incompatibility:

‘ $s \models \sim A$ ’ iff for any x if $x \models I A$ then Nsx

¹⁶ Pg 81 Mare

¹⁷ I am assuming a non-dialethetist para-consistency framework...

Definition 2.9 Negation in terms of compatibility:

‘ $s \vDash \sim A$ ’ iff for any x if Csx then it is not the case that $x \vDash A$

We will follow Mares (2004) in primarily using the compatibility relation to define a relevant negation. Since we have inconsistent situations, this precludes the compatibility relation from being reflexive. However in relevant logic, $A \rightarrow \sim\sim A$ is a rule in its natural deduction system. This requires that the compatibility relation be symmetrical, that is if Cst then Cts . This shouldn’t be surprising considering that the structural rule of weak commutativity ($X;Y \vdash Y;X$) is included in the relevant sequent calculus.

A Short Tangent: An additional negation operator

Symmetrical compatibility will lead us astray in formalizing deontic logic, as we will discover later in this chapter. Thus, we will briefly investigate how we may introduce an additional negation operator by rejecting both versions of commutativity and introducing a right to left conditional operator. Recall that in the sequent calculus where A, B and X are structures and ‘;’ is our punctuation we have our left to right arrow introduction rule¹⁸ (\rightarrow I):

$$\frac{X; A \vdash B}{X \vdash A \rightarrow B}$$

For a right to left arrow introduction rule (\leftarrow I) we have:

$$\frac{A; X \vdash B}{X \vdash B \leftarrow A}$$

With weak commutativity¹⁹, \rightarrow and \leftarrow are equivalent. However, to achieve a good formalization of reasoning about obligation and its cognates, I propose that we need to have both arrows. Briefly and informally, the need for both arrows in this domain is due

¹⁸ Which is predicated upon our semantic entailment version of the deduction theorem

¹⁹ Among other substructural logics, there are other ways of making \rightarrow and \leftarrow equivalent. For example, if we have 0 as a zero-place punctuation mark for the empty-set, we achieve the same result with commutativity and the push and pop rules.

to how we seem to be most concerned about obligations regarding actions. And reasoning about actions appears to require both arrows, especially when reasoning about one action precluding another action: one action may be incompatible with the other, but not the other way around. For example, Restall appeals to the pair <giving all your money away, buying a sports car>, which we will symbolize as M and C respectively. This pair of actions is jointly incompatible, but in only one direction:

$$\frac{M; C \vdash f}{M \vdash C \rightarrow f}$$

The above sequent, together with our definition of ‘ \sim ’ negation, provides us with $M \vdash \sim C$; but it seems undesirable to be able to prove $C \vdash \sim M$ from $M \vdash \sim C$ in this case. However, with weak commutativity it is so provable:

$$\begin{array}{ll} \frac{M \vdash \sim C}{M; C \vdash f} & \text{(hypothesis)} \\ \frac{M; C \vdash f}{C; M \vdash f} & \text{(by } \sim \text{ definition and } \rightarrow\text{e)} \\ \frac{C; M \vdash f}{C \vdash M \rightarrow f} & \text{(weak commutativity)} \\ & (\rightarrow\text{I}) \end{array}$$

In words, giving away all my money precludes then buying a sports car. But buying a sports car should not preclude then giving away all my money (even if all that’s left over is the change in the sofa). Therefore, by rejecting commutativity and by including the right to left arrow, we obtain a different inference to f :

$$\frac{M; C \vdash f}{C \vdash f \leftarrow M}$$

We may then use $f \leftarrow A$ to define a distinct negation:

Definition 2.10 \neg negation:

$$\neg A = f \leftarrow A$$

Thus we may interpret $M \vdash \sim C$ as claiming that giving all your money away precludes you from also buying a car. We may interpret $C \vdash \neg M$ as claiming that you buy a sports car, given that you didn’t first give all your money away. In relevant logic

the order of the premises doesn't normally matter, but in order to formalize reasoning regarding actions, the order of the actions appears to be relevant. Since ethical reasoning is for the most part concerned about obligations regarding actions, it is likely that the order of premises is relevant for characterizing the appropriate inference relation of deontic logic. Moreover, in the same manner that one action may preclude the other, but not *vice versa*, we may need the same sort of structural rules to deal with one obligation trumping another, since trumping should not, likewise, be a symmetrical relation.

Negation and Partial and Inconsistent Situations

The relevant ' \sim ' negation as outlined above, as defined by the compatibility relation, indicates constraints on the properties that the points in our models may possess. Consider a situation s , for example the situation that "consists of the information that is currently available to me...[n]othing happening here makes it true that it is currently raining ... on the other side of the globe. But situations in which it is raining [on the other side of the globe] are compatible with my current situation. So neither 'It is raining [on the other side of the globe]' nor 'It is not raining [on the other side of the globe]' is true in my current situation."²⁰ Therefore neither R (for raining etc) nor $\sim R$ is true of situation s . The situation s is thus partial.

Likewise, consider another situation s , which is someone's representation of a part of the world. Situation s however represents its corresponding part of the world in incompatible ways. Situations like s would then be incompatible with themselves. Mares argues that "[w]hen a situation s is incompatible with itself, it is possible for it to make a formula A true, but A fail to be true in every situation compatible with s ."²¹

²⁰ Pg 75 Mares

²¹ Pg 76 *ibid*

So: s belongs to $I(A)$ but for any x if Csx then it is not the case that $x \models A$. Therefore $s \models I A \& \sim A$.

Box and Diamond Operators for R

Since in this chapter we have used R to characterize the accessibility relation for implication, I will use M for a modal accessibility relation, to avoid confusion. We can use a familiar take on standard kripkean semantics to define our box and diamond operators:

Definition 2.11 A relevant \Box :

$$'s \models I \Box A' \text{ iff } \forall x, \text{ if } Msx \text{ then } x \models I A$$

Definition 2.12 A relevant \Diamond :

$$'s \models I \Diamond A' \text{ iff } \exists x \text{ such that } Msx \& x \models I A^{22}$$

Given an application of these modal operators to deontic logic, we have a revised, but familiar, interpretation of ought and permitted:

it is true that ' A is obligatory' in a situation s , if and only if, A is true at every morally ideal situation accessible from s .

it is true that ' A is permitted' in a situation s , if and only if, A is true in at least one morally ideal situation accessible from s .

Moreover, we will need the D axiom ' $\Box A \rightarrow \Diamond A$ ' as an axiom in relevant deontic logic, much like in standard deontic logic. The semantic postulate corresponding to this axiom

²² Note that with commutativity, symmetrical compatibility, and with a postulate regarding the existence of a maximal consistent situation x for every situation s , such that Csx , and for any y if Csy then x extends y , we then have double negation elimination (dne) as a rule, which allows us to instead define \Diamond as $\sim\Box\sim$. However, I happen to be attracted to the adoption of an intuitionistic version of negation, since the semantic postulates in relevant logic required for dne are cumbersome and require too much faith in the appeal of dne as a good inference.

is again identical to the one in SDL: that for any s there is an x such that Msx . It seems inappropriate to justify an obligation to ‘A’ due to ‘A’ being true at all morally accessible situations, just because there are no accessible situations.

Taking Genuine Moral Dilemmas Seriously

With the logic $R\Box$ for relevant modal implication, impossible moral situations can take a fruitful part in reasoning about obligation and its cognates. Recall that by a ‘genuine moral dilemma’, we refer to situations in which incompatible moral demands confront a moral agent, without any reason to select one obligation as having more weight or authority than the other. By ‘fruitful’, I mean that valid reasoning from an inconsistent set of premises does not remove all constraints on what can be inferred from that set of premises. If it is possible to fruitfully reason from an inconsistent set of premises then I argue that such a type of reasoning deserves to be considered as taking genuine moral dilemmas seriously.

Theorem 2.1 $R\Box$ does not remove all constraints on what can be inferred from an impossible moral situation of type ‘ $\Box A \ \& \ \Box \sim A$ ’:

The inference $\Box P \ \& \ \Box \sim P$ therefore B; for arbitrary B, is invalid in $R\Box$

Proof Suppose we have $s \Vdash \Box P \ \& \ \Box \sim P$, then $\forall x$, if Msx then $x \Vdash P$ and $x \Vdash \sim P$. From D, we infer there is an x , such that $x \Vdash P \ \& \ \sim P$. By definition of ‘ \sim ’ negation, we have that for any y , if Cxy then it is not the case that $y \Vdash P$. However there is nothing in this model that forces that $x \Vdash B \vee \sim B$, since x may be partial. Thus while $\forall x$, such that Msx , $x \Vdash P \ \& \ \sim P$, we do not have $x \Vdash B$ for arbitrary formula B. Furthermore, $\forall y$ such that $Rsxy$, there is nothing in the semantics that forces $y \Vdash B$, so $(\Box P \ \& \ \Box \sim P) \rightarrow B$ is also not a valid formula of $R\Box$.

Theorem 2.2 $\mathbf{R}\Box$ does not remove all limits on what can be inferred from an impossible moral situation of type ‘ $\Box A \ \& \ \sim\Box A$ ’:

The inference $\Box P \ \& \ \sim\Box P$ therefore B; for arbitrary B, is invalid in $\mathbf{R}\Box$

Proof Suppose we have $s \Vdash \Box P \ \& \ \sim\Box P$. Then by \Box definition we have that for any x if Msx , then $x \Vdash P$. But we also have that it is the case that not every x , such that Msx , $x \Vdash P$. So we have an x , in which $x \Vdash P$ and $x \Vdash \sim P$. So every situation y such Cxy , y can not make P true. However, since $x \Vdash P$, Nxx . However, there is nothing in the semantics that forces $x \Vdash B \vee \sim B$, since x may be partial. Furthermore, for any y such that $Rsxy$, there is nothing in the semantics that forces $y \Vdash B$, so we also do not have $(\Box P \ \& \ \sim\Box P) \rightarrow B$ as a valid formula in $\mathbf{R}\Box$.

Theorem 2.3 $\mathbf{R}\Box$ does not remove all limits on what can be inferred from an impossible moral situation of type ‘ $\Box A; \ \Box B \vdash f$ ’:

$(\Box A \ \& \ (\Box B \rightarrow f)) \rightarrow C$; for arbitrary C, is invalid in $\mathbf{R}\Box$

Proof Suppose we have an arbitrary situation s , such that $Rstu \ \&$ for any situation y , if $y \Vdash \Box A \ \& \ \Box B$, then it is also the case that $y \Vdash f$. Let’s assume that $t \Vdash \Box A \ \& \ \Box B$. Then we also have $t \Vdash f$. Since u extends t , $u \Vdash f$. Thus we have $s \Vdash \sim(\Box A \ \& \ \Box B)$ and $t \Vdash \Box A \ \& \ \Box B$; therefore Nst . However, there is nothing in the semantics that forces $u \Vdash C$ since u may be a partial as well as an inconsistent situation. **Q.E.D.**

These counter-examples to obligation explosion in $\mathbf{R}\Box$ may be then cashed out as the class of models which provides value assignments in which all accessible morally ideal situations are inconsistent but partial situations.

Formalizing Trumping Hierarchies

When formalizing a trumping hierarchy for ethical reasoning, we desire a system that can capture the notion that even though something may be true in every accessible morally ideal situation, it may be over-ridden by some other proposition true in every accessible morally ideal situation. In SDL, we discovered we were committed to there only being one member of a pairwise inconsistent set being true at a time in any morally ideal world. As such it didn't make sense to say that one obligation trumped another obligation, since obligations were defined as being true at all accessible morally ideal situations. And if only one of two inconsistent propositions can be true in any accessible morally ideal world, only one of those propositions could be seen as being obligatory, since at most one can be true in every such world. Given that it appears infeasible or undesirable to jettison from ethical theories all talk²³ of one obligation over-riding another; it seems that SDL doesn't deserve to be considered a good candidate for deontic logic.

An appropriate formalization of trumping-like reasoning must be able to let obligations that are trumped by another obligation, remain obligations nonetheless, i.e., they would remain true at all accessible morally ideal situations. Thus the semantics of trumping seemingly require inconsistent points in its framework. Relevant logic happens to have such a framework. To formalize trumping in a hierarchy we introduce a binary operator $<$ on formulae, which provides a partial ordering. Thus, if $P < Q$, we infer that Q should have more moral weight in our moral reasoning than P .

²³ Granted that this talk takes obligations as more than merely hypotheses entered into a chain of reasoning

Theorem 2.4 $R\Box$ is compatible with the claim that, if $P < Q$; i.e., assuming $Q; P \vdash f^{24}$, then it is obligatory to not do P, even though P is true in all accessible morally ideal situation.:

$$\Box P \rightarrow (\Box Q \rightarrow (\Box P \ \& \ \Box \sim P))$$

Proof Suppose we have $s \not\vdash I \Box P$. Then, for any x such that Msx then $x \not\vdash I P$. Let's then also suppose that $s \not\vdash I \Box Q$, thus for any x such that Msx then $x \not\vdash I Q$. According to our assumption $Q; P \vdash f$, we have by the semantic entailment theorem, $x \not\vdash I P \rightarrow f$, and by our definition of \sim negation, we have $x \not\vdash I \sim P$. Since for any x , Msx that $x \not\vdash \sim P$, we have that $s \not\vdash \Box \sim P$. Nonetheless, we still have (since every accessible x is an inconsistent point) that every x , such that Msx , $x \not\vdash I P$; so we also have $s \not\vdash I \Box P$ as desired.

Theorem 2.5 Unfortunately, similar reasoning shows that with a hierarchy containing that $P < Q$ and assuming $Q; P \vdash f$, we cannot block the following undesirable feature of $R\Box$. Due to weak commutativity and symmetrical compatibility being a legitimate feature of R :

$$\Box P \rightarrow (\Box Q \rightarrow (\Box \sim Q \ \& \ \Box \sim P))$$

We do not want this to be a valid scheme, given that we want a hierarchy with $P < Q$ to mean that Q precludes P, not that P also precludes Q²⁵

Proof Suppose we have $s \not\vdash I \Box P$, then for any x such that Msx then $x \not\vdash I P$. Let's then suppose that $s \not\vdash I \Box Q$, thus for any x such that Msx then $x \not\vdash I Q$. According to our assumption $Q; P \vdash f$, we have by the semantic entailment theorem, $x \not\vdash I P \rightarrow f$, and by our definition of \sim negation, we have $x \not\vdash I \sim P$. By \Box -definition, we infer that $s \not\vdash I \Box \sim P$. By commutativity, $Q; P \vdash P; Q$ and thus $P; Q \vdash f$. Therefore at x we also have $x \not\vdash I Q \rightarrow f$, and

²⁴ That is, given any arbitrary s such that if $s \not\vdash P \ \& \ Q$, then $s \not\vdash f$, then for any x , for any y such that $Rsxy$ if $x \not\vdash P$ and $y \not\vdash Q$ then Nxy

²⁵ Unless there is also a competing hierarchy, where $Q < P$...

by our definition of \sim negation, we have $x \Vdash \sim Q$. From \Box -definition, we infer that $s \Vdash \Box \sim Q$ which was to be proved.

Proposition 2.0 The attentive reader has probably already guessed at my proposed remedy: With a hierarchy containing $P < Q$ when $Q; P \vdash f$, and rejecting (weak) commutativity and using both \leftarrow and \rightarrow , we can formalize a trumping function that returns the following desirable result:

$$\Box P \rightarrow (\Box Q \rightarrow (\Box \sim P \ \& \ \sim \Diamond \sim Q))$$

I take this formula to represent the claim that when Q trumps P , even though Q is obligatory and P is obligatory, we are obligated to not P when we must choose between Q and P . However, it is not the case that we are also permitted or obligated to not Q : it is not a matter of moral indifference which one we choose to do.

Proof Suppose we have $s \Vdash \Box P$, then for any x such that Msx then $x \Vdash P$. Let's then suppose that $s \Vdash \Box Q$, thus for any x such that Msx then $x \Vdash Q$. According to our assumption $Q; P \vdash f$, we have by the semantic entailment theorem, $x \Vdash P \rightarrow f$, and by our definition of \sim negation, we have $x \Vdash \sim P$. By \Box -definition, we infer that $s \Vdash \Box \sim P$. Rejecting commutativity, and with $\leftarrow I$, we also have $x \Vdash f \leftarrow Q$ and by our definition of \neg , $x \Vdash \neg Q$. So examining every x such that Msx , we find that $x \Vdash P \ \& \ Q \ \& \ \sim P \ \& \ \neg Q$, but there is not a single x , such that $x \Vdash \sim Q$ as desired. **Q.E.D.**

The Naturalistic Fallacy

In the semantics of $R\Box$, as outlined above, we distinguish between logical situations and non-logical situations in order to partly justify and interpret a ternary accessibility relation R for relevant implication. However, the distinction between situations which are complete and consistent (logical) and those which are either partial

or inconsistent (or both) will help also to formalize the claim that it is possible that from some point s there are no obligations.

The culprit in SDL being incapable of formalizing the possibility of no obligations claim, is that it is closed under necessitation, i.e.,

$$\frac{\vdash A}{\vdash \Box A}$$

With relevant models being populated with both logical and non-logical situations, we notice the following semantic fact:

$$\frac{s \not\vdash I A}{s \not\vdash I \Box A} \quad \text{iff } s \text{ is logical and for any } t, Mst \text{ then } t \text{ is logical.}$$

Consider a partial (and thus non-logical) t , i.e., a situation that for some A , it makes neither A nor $\sim A$ true. Then even if $\vdash A$, and Mst (assuming that s is logical, thus $s \vdash I A$), nonetheless, it is not the case that every x such that Msx , $x \vdash I A$ as required by the definition of \Box : formulae verified by the empty-set can not be automatically verified at non-logical situations. The situation t , according to our assumption, is such a situation: it lacks the information that $\vdash A$.

Moreover, suppose that we do have $\vdash \Box A$; we do not have to worry about being able to infer $\vdash p \rightarrow \Box A$ from $\vdash \Box A$ in $\mathbf{R}\Box$. Weakening is not a rule in \mathbf{R} .

Summary

While $\mathbf{R}\Box$ seemed able to formalize using genuine moral dilemmas in ethical reasoning, and dodged making the naturalistic fallacy valid, there were some problems with trumping due to the symmetry of the compatibility relation. Moreover, the order of the premise structures in the sequent calculus must be identical to the order in the trumping hierarchy. However, it is not immediately clear that such a fit is automatically

justified, as we may not be able to switch back and forth from the incompatibility of actions and the theories about which moral demands carry more weight. One may have very little to do with the other.

For example, suppose I am a pious Christian, given to taking the Bible literally. Imagine then, that I read²⁶ how Jesus said to give all your possessions away to the poor; and since no pious Christian should believe that it is possible for any morally ideal world to not make the words of Jesus true, I recognize an obligation to give all my money to the poor. However, if I give all my money away I will be unable to afford the expensive medical treatment required to keep my child healthy. Thus giving all my money away precludes me from then keeping my child healthy; however, consider a hierarchy which posits that the responsibility to care for one's dependents should trump giving all your money away to strangers:

$M; H \vdash f$ and $M < H$ but $s \not\models \Box M$ and $s \models \Box H$, where M stands for 'giving all your possessions to the poor' and H stands for 'caring for your sick child'.

What we are looking for is to be able to infer that it is permissible to not give away all your money, given that you have a sick child to care for.

Suppose we have $s \models \Box H$, then for any x such that Msx then $x \models H$. Let's then also suppose that $s \models \Box M$; thus for any x such that Msx then $x \models M$. According to our assumption $M; H \vdash f$, we have by the semantic entailment theorem, $x \models H \rightarrow f$, and by our definition of \sim negation, we have $x \not\models \sim H$. From D-schema, we infer that there is an $x \models \sim H$, and thus by the definition of \diamond , we have $s \models \diamond \sim H$. And this is not at all desirable: we may interpret this as claiming that it is permissible to not care for your sick child's health.

²⁶ Begging off questions on appropriate exegesis

What we do want is that $M \vdash \neg H$, which is meant to formalize that we can deduce from someone caring for one's child that the one did not first give away all their money. With a deontic interpretation, it may be represent the claim that I can discharge an obligation to give all my possessions away, provided that I don't have a sick child I ought to care for.

Therefore, we need a semantic postulate governing the interaction of a hierarchical ordering and the order of premise structures in the sequent calculus; moreover, the hierarchical ordering must be able to over-ride the action ordering. If we can't make sense of the hierarchical ordering replacing the action ordering, it seems that there is no way to block the inference (that is, without jettisoning the deduction theorem) that, since giving all your money away precludes caring for your child's health, it is then permissible to not care for your child's health if you are obligated to give all your money away. And this is incompatible with the basic notion behind the trumping approach to resolving conflicting moral demands: that it is not due to merely which action precludes which action, but that it should be the hierarchy which alone determines which action trumps which action. In appendix A, I sketch out a preliminary attempt at cashing out how a hierarchical ordering could replace the prior action ordering.

Thus by considering the incompatibility of actions at accessible morally ideal situations, we end up with a poor approximation of which actions are permissible or impermissible. Yet it is the very incompatibility of actions which enables impossible moral situations. Therefore, there seems to be considerable tension with simply using $\sim A$ being true at some accessible morally ideal situation as defining 'it is permissible that $\sim A$ '.

Chapter Three

Multiple-Accessibility Relational Structures

So far, we have been largely concerned with the relationship between the concepts of obligation and permissibility or possibility. However, Peter Schotch and Raymond Jennings have argued that a viable formalization of deontic reasoning should instead focus on jettisoning the commonly accepted *aggregation* principle they refer to as the K-schema: $(\Box P \ \& \ \Box Q) \rightarrow \Box(P \ \& \ Q)$. In chapter one, we identified a different schema as the K-schema, namely: $\Box(P \rightarrow Q) \rightarrow (\Box P \rightarrow \Box Q)$. Nonetheless these two K-schemas are equivalent in all normal frames, i.e., frames for which ‘ $\vdash a \rightarrow b$ therefore $\vdash \Box a \rightarrow \Box b$ ’ and necessitation holds. For the purposes of this chapter, we will be using $(\Box P \ \& \ \Box Q) \rightarrow \Box(P \ \& \ Q)$ as our K-schema, keeping in line with Schotch and Jennings usage.

By ‘aggregation principle’ Schotch and Jennings mean any principle which claims that “any finite number of necessities can be aggregated to produce one.”²⁷ The issue that Schotch and Jennings take with the aggregation principle expressed by the K-schema, is that it is too strong: “it collapses deontically significant distinction between modal sentences”²⁸ They argue that the D-schema, that is, $\Box P \rightarrow \Diamond P$, and a schema representing that there are no obligations to bring about an impossible state of affairs, that is, $\sim \Box f$, must be kept logically distinct in any viable formalization of deontic reasoning. However, even the weakest modal logic that makes the K-schema valid, will conflate these two schemas.

Proof: Let us suppose that for some x , $x \models \Box P$ and $x \models \sim \Diamond P$, i.e. D-schema fails. Thus $x \models \Box P$ and $x \models \Box \sim P$. From K-schema we infer that $x \models \Box(P \ \& \ \sim P)$; of course, $P \ \& \ \sim P \vdash$

²⁷ Pg 152

²⁸ ibid

f , and in virtue of $\sim\Box f$, this result demonstrates that we can't have $\sim\Box f$ without making D-schema valid. Proving the equivalence in the other direction is trivial.

Schotch and Jennings want to rule out 'it is obligatory to do the logically absurd', but keep open the possibility that there are conflicting obligations, i.e., those situation when we are obligated to A and we are obligated to not A. Thus, the way forward seems to require theorists to reject principles which aggregate distinct obligations into one comprehensive obligation. We will therefore examine Schotch and Jennings' proposal to block the inference from an obligation to do 'A' and an obligation to do 'B', to there being an obligation to do both 'A and B'. We will also investigate whether we can make any headway in formalizing trumping-like reasoning, and whether Schotch and Jennings' proposal provides a formalization which doesn't validate inferring an 'ought' from an 'is' or ruling out the possibility that there are or were no obligations from some point.

Schotch and Jennings' Multiply Accessibility Relations Semantics

It seems that, at least *prima facie*, many moral agents are committed to several evaluative frameworks at once, rather than one monolithic moral system. Ordinary day-to-day ethical reasoning is seemingly characterized by agents switching back and forth, sometimes quite unreflectively, from harm-avoidance consequentialist theories of right and wrong to duty theories to theories of liberty, even when dealing with the evaluation of a single moral situation.

Consider the common response one may hear when many people find out that a lawyer is defending an 'obviously' guilty person: how can she, the lawyer, try to get this person back on the streets without having to pay for what he or she did? The lawyer may agree that while any morally ideal situation is such that her guilty client is locked up,

nonetheless she also has a duty to provide the best defense, within reasonable limits, for her client. Perhaps her first evaluation may be cashed out as a harm-avoidance claim, while the second is best understood as grounded in claims regarding moral limits to what can be required of the individual for the sake of the common good. In this case, the lawyer is claiming that the second evaluation provides reasons to trump the results of the former evaluation, but we will get to that later.

Moreover, it is common practice, and rightly so, to think that an agent is morally deficient in some fashion when he or she utterly disregards the values underlying cogent evaluative schemes; especially so, when it just so happens that these other evaluative schemes provide evaluations which conflict with their favorite theory. It is not merely that such people are being unreasonable for not considering other evaluative frameworks, but when anyone rejects the basic values underlying any cogent evaluative framework, this reflects poorly on his or her own moral status.

To enter unabashedly into controversial territory, consider the abortion debate. Neither side seems willing to concede that the opposing side is grounded (however weakly) on a basic claim of value: that every case of a specific collection of biological cells organized in such and such way deserves to be considered of intrinsic moral worth and that there are moral limits to what can be required of any individual for the sake of the good of others²⁹. Unfortunately, in this often heated debate, proponents of one side frequently dismiss the basic values underlying the other sides' position, resulting not only in an all-round unfruitful discussion, but also in people becoming committed to immoral claims.

²⁹ Moreover, both sides appeal to both of these claims, in some form, while somehow denying they apply to the other position.

It thus seemingly behooves theorists to consider whether different theories or evaluative schemes pick out different sets of morally ideal situations. For Schotch and Jennings, the possible worlds of classical modal logic are sufficient for characterizing the appropriate framework of deontic logic. Therefore, given a non-empty set of possible worlds, Schotch and Jennings propose that we allow for “two or more notions of accessibility to operate simultaneously” in picking out the set of possible worlds we use to define moral necessity or equivalently, obligation.

Definition 3.0 A frame is a relational structure F such that $F = \langle W, R_1, \dots, R_n \rangle$, where W is a non-empty set of possible worlds and every $R_i (1 \leq i \leq n)$ is a binary relation between possible worlds of W .

Definition 3.1 A model is a relational structure M such that $M = \langle W, R_1, \dots, R_n, I \rangle$ where $\langle W, R_1, \dots, R_n \rangle$ is a frame and I is an interpretation which associates with every atomic sentence of the language a set of possible worlds of W in which the sentence is true.

Definition 3.2 Moral necessity at a world: ‘ $w \models I \Box A$ ’ is true iff there is an $R_i (1 \leq i \leq n)$ such that $\forall x$ (if $R_i wx$ then $x \models I A$).

Definition 3.3 Moral permissibility at a world: ‘ $w \models I \Diamond A$ ’ is true iff $w \models I \sim \Box \sim A$

Proposition 3.0 The K-schema is not valid in the above class of relational structures.

Proof: We need a model such that for some w , $w \models I \Box A$, $w \models I \Box B$ and $w \models I \sim \Box (A \& B)$.

Thus we need that there is some $R_i (1 \leq i \leq n)$ such that $\forall x$ if $R_i wx$ then $x \models I A$, and that there is some $R_j (1 \leq j \leq n)$ such that $\forall x$ if $R_j wx$ then $x \models I B$. Let $i = 1$ in the first case and $i = 2$ in the second. Suppose that $R_1 = \{ \langle w, y \rangle \}$ and that $R_2 = \{ \langle w, z \rangle \}$. Let $y \models I A$ and $y \models I \sim B$ and let $z \models I B$ and $z \models I \sim A$, and that there are only the three worlds w , y , and z .

Thus $\forall x$ if $R_1 wx$ then $x \not\models A$ and $\forall x$ if $R_2 wx$ then $x \not\models B$, but there is no i ($1 \leq i \leq n$) such that $\forall x$ if $R_i wx$ then $x \not\models A \& B$.

We now turn to clarifying how the above class of relational structures will nonetheless validate $\sim \Box f$. Similar to the *seriality* property of SDL, we postulate the following existence condition:

Definition 3.3 Existence condition for Multiple Relations Semantics:

$$\forall x \exists y : R_1 xy \text{ or } \dots \text{ or } R_n xy$$

Proposition 3.1 The above class of relational structures, together with the existence condition, validates $\sim \Box f$ without validating $\Box P \rightarrow \Diamond P$

Proof: Using classical possible worlds, there is no world x , such that $x \not\models f$. Therefore there is always an x such that $R_i wx$ and $x \not\models f$. However, suppose we use the same three worlds from above, with $R_1 = \{ \langle w, y \rangle \}$ and that $R_2 = \{ \langle w, z \rangle \}$. Let us assume that $y \not\models A$ and $z \not\models \sim A$. This means that $w \not\models \Box A$ but $w \not\models \sim \Diamond A$, as desired.

Genuine Moral Dilemmas

Schotch and Jennings' deontic logic (SJDL) does not remove all constraints from what can be inferred from an inconsistent or incompatible set of obligations.

Proposition 3.2 SJDL does not remove all constraints on what can be inferred from an impossible moral situation of type ' $w \not\models \Box A, \Box \sim A$ ':

The inference $\Box P, \Box \sim P$ therefore Q ; for arbitrary Q , is invalid in SJDL

Proof: Suppose again that there are but the three worlds as above and that $R_1 = \{ \langle w, y \rangle \}$ and that $R_2 = \{ \langle w, z \rangle \}$. Suppose also that $y \not\models A$ and that $z \not\models \sim A$, thus $w \not\models \Box A, \Box \sim A$, but nothing in the semantics forces $w \not\models Q$ or $y \not\models Q$ or $z \not\models Q$, for some arbitrary Q .

Proposition 3.3 SJDL does not remove all limits on what can be inferred from an impossible moral situation of type ‘ $w \vDash I \Box A, \sim \Box A$ ’:

The inference $\Box P, \sim \Box P$ therefore Q ; for arbitrary Q , is invalid in SJDL

Proof: Suppose that there are but the worlds w, y, z_1 and z_2 , and that for any i for $R_i, i = 1$ or $i = 2$. We need a model such that $R_1 = \{ \langle w, y \rangle \}$ and that $R_2 = \{ \langle w, z_1 \rangle, \langle w, z_2 \rangle \}$.

We also suppose that $y \vDash I A$ and that $z_1 \vDash I A$ and $z_2 \vDash I \sim A$. Therefore $w \vDash I \Box A, \sim \Box A$.

However nothing in the semantics forces $w \vDash I Q$ or $y \vDash I Q$ or $z_1 \vDash I z Q$ or $z_2 \vDash I Q$ for some arbitrary Q .

Proposition 3.4 SJDL does not remove all limits on what can be inferred from an impossible moral situation of type ‘ $w \vDash I \Box A, \Box B$ and $A \& B \vdash f$ ’:

The inference $\Box P, \Box Q, P \& Q \vdash f$, therefore C ; for arbitrary C , is invalid in SJDL

Proof: For our usual three worlds, suppose that $R_1 = \{ \langle w, y \rangle \}$ and that $R_2 = \{ \langle w, z \rangle \}$.

Suppose also that $y \vDash I A$. thus $y \vDash I \sim B$ since y is a classical possible world and $A \& B \vdash f$.

Suppose also that that $z \vDash I B$, and $z \vDash I \sim A$ as above. Thus $w \vDash I \Box A, w \vDash I \Box B$. However, nothing in the semantics forces that $w \vDash I C$ or $y \vDash I C$ or $z \vDash I C$, for some arbitrary C .

Formalizing Trumping

As noted in chapter two, we desire a system that can formalize the notion that even though something may be true in every accessible morally ideal situation, it may be over-ridden by some other proposition true in every accessible morally ideal situation. However, with classical worlds, we are committed to only one member of an inconsistent pair being true at a time in any morally ideal world. But with multiple accessibility relations in SJDL, perhaps progress can be made towards formalizing how an obligation

verified by one accessibility relation may be trumped by another obligation verified by a different accessibility relation.

Proposition 3.5 SJDL is compatible with the claim that when $Q; P \not\vdash f$, but there is a hierarchy such that $P < Q$, then it is obligatory and permissible to not do P, even though P is true in all morally ideal worlds accessible with some R-relation. That is, it is compatible with the truth of:

$$\Box P \rightarrow (\Box Q \rightarrow (\Box P \ \& \ \Box \sim P))$$

Proof: Suppose that $W = \{x, y, z\}$ and that $R_1 = \{< w, y >\}$ and that $R_2 = \{< w, z >\}$.

Suppose also that $y \not\vdash I Q$. Then $y \not\vdash I \sim P$, since y is a classical possible world and $P \ \& \ Q \not\vdash f$.

Suppose also that $z \not\vdash I P$. Thus $w \not\vdash I \Box Q$, $w \not\vdash I \Box P$ and $w \not\vdash I \Box \sim P$, as desired.

Proposition 3.6 Unfortunately, similar reasoning shows that with a hierarchy containing that $P < Q$ and assuming $Q; P \not\vdash f$, we cannot block the following undesirable feature of SJDL:

$$\Box P \rightarrow (\Box Q \rightarrow (\Box \sim Q \ \& \ \Box \sim P))$$

We do not want this to be a valid scheme, given that we want a hierarchy with $P < Q$ to mean that Q precludes P, not that P also precludes Q.

Proof: Suppose that $W = \{x, y, z\}$ and that $R_1 = \{< w, y >\}$ and that $R_2 = \{< w, z >\}$.

Suppose also that $y \not\vdash I Q$, therefore $y \not\vdash I \sim P$ since y is a classical possible world and $P \ \& \ Q \not\vdash f$.

Suppose also that $z \not\vdash I P$, therefore $z \not\vdash I \sim Q$ as above. Thus $w \not\vdash I \Box Q$, $w \not\vdash I \Box P$. We

also have that $w \not\vdash I \Box \sim P$, and $w \not\vdash I \Box \sim Q$, which was to be shown

I propose that, once again, in order to formalize trumping hierarchies, we need to postulate inconsistent situations

Proposition 3.7 With inconsistent situations and multiple accessibility relations we can formalize trumping-like reasoning. What we want when $P < Q$ and $Q; P \Vdash f$ is for this to be true:

$$\Box P \rightarrow (\Box Q \rightarrow (\Box \sim P \ \& \ \sim \Diamond \sim Q))$$

Proof: Suppose that $W = \{x, y, z\}$ and that $R_1 = \{\langle w, y \rangle\}$ and that $R_2 = \{\langle w, z \rangle\}$.

Suppose also that $y \Vdash P$ and $y \Vdash Q$. (We can suppose this by taking y to be an inconsistent situation instead of a classical possible world.) Suppose, on the other hand, that that $z \Vdash Q$, but that $z \Vdash \sim P$. Thus we have $w \Vdash \Box P$, $w \Vdash \Box Q$, $w \Vdash \Box \sim P$ but not that $w \Vdash \Diamond \sim Q$, as desired.

However, with inconsistent situations, we would now have to reconsider how we characterize our semantics so as to avoid making $w \Vdash \Box f$ possible. The non-logical situations in this revised semantics are already inconsistent; it seems that we can only gain by allowing non-logical situations to also be incomplete. However, to avoid verifying obligations to do the impossible, we first must preclude the possibility that our non-logical situations can simply make ‘ f ’ true. Secondly, to rule out cases like ‘ $w \Vdash \Box(A \ \& \ \sim A)$ ’, the non-logical situations in this revised semantics must also not be closed under other classical logical laws, in this case, *aggregation*:

Semantic Postulate: Trumping SJDL will not verify the following:

For any x such that $x \in \text{non-logical}$, if $x \Vdash A$ and $x \Vdash B$ therefore $x \Vdash A \ \& \ B$.

For any x such that $x \in \text{non-logical}$, $x \Vdash f$.

Therefore, in the proof for Prop 3.7, while $y \Vdash P$ and $y \Vdash Q$, since $y \in \text{non-logical}$, there is nothing in this semantics which forces that $y \Vdash (P \ \& \ Q)$. Therefore we do not have $w \Vdash \Box(P \ \& \ Q)$ as desired.

These revisions shouldn't be as troubling as it first appears; non-logical worlds are already not closed under explosion, so we might have to make peace with the possibility that, in order to formalize trumping-like reasoning, few familiar classical laws of logic can apply to morally ideal situations. Thus we justify our semantic rules because they characterize what we take as cogent reasoning in our shared experiences of moral inquiry, not because the rules are merely familiar and simple to use.

The Naturalistic Fallacy

Regardless of how many accessibility relations we have in our semantics, as long as it is only consistent and complete points which populate our relational structures, if the empty set forces a formula, then every point make that formula true. Therefore, we have:

$$\frac{\vdash A}{\vdash \Box A}$$

For any n of R_n ($1 \leq n$) and for any x for some w , such that $R_n wx$ then $x \Vdash A$, since in a classical model, every x , $x \in W x \Vdash A$.

If we have $\Vdash A$, then we can infer that $P \Vdash A$, since weakening with the empty set is still a valid rule in SJDL: the \rightarrow is still understood as the horseshoe. That is, if A is true at every point, then $P \rightarrow A$ is also true at every point. Furthermore, if we can infer that A from the empty set, we can infer that A from any P as well in classical proof-theory.

Thus, we have the following undesirable feature in SJDL:

$$\frac{\frac{\frac{\vdash A \text{ (hypothesis)}}{\vdash \Box A \text{ (closure under } \Box)}}{p \vdash \Box A \text{ (weakening)}}}{\vdash p \rightarrow \Box A \text{ } (\rightarrow I)}$$

If we interpret the operators with deontic terms, we have the inference of an 'ought' from an 'is', as being valid.

Schotch and Jennings claim that necessitation closure is not a “central principle. Its function is to allow us to employ a more streamlined semantics than would otherwise be possible. Should [this rule] be thought genuinely counter-intuitive, rather than just odd, our subscription to it may be terminated. The semantics of the resulting logic(s) must then be complicated along lines familiar from the analysis of so-called non-normal modal logics.”³⁰ The complicated semantics are as follows:

We leave behind classical possible worlds semantics for the semantics of non-logical situations. Fortunately, we have already justified the use of inconsistent situations, and situations that are not closed under *aggregation*; we can only gain, as I argued above, by using incomplete situations as well. Thus, since incomplete situations can fail to have the information that $\vdash A$, while $\vdash A$ necessarily holds for all logical situations, we cannot verify the following:

‘ $\vdash A$ and for any n of R_n ($1 \leq n$) and for any x for some w , such that if $R_n wx$ then $x \not\vdash A$ ’ unless w and x are logical. Thus, closure under necessitation fails in this semantics, since we have non-logical situations.

Summary

The system proposed by Schotch and Jennings does provide for the formalization of fruitful reasoning from inconsistent or incompatible sets of obligations. However, we discovered that in order to formalize trumping and ruling out deriving ‘ought’ from ‘is’, classical possible worlds are insufficient for the semantics of SJDL.

Once we begin quantifying over such a larger class of interpreted objects, we found that SJDL quite simply formalized trumping reasoning by using another

³⁰ Pg 159 S&J

accessibility relation to pick out a different set of morally ideal situations. With at least two sets of morally ideal situations, we can provide the models required by reasoning in which one obligation trumps another obligation. Thus we have that both obligations remain true at all morally ideal situations accessible with some morally accessibility relation, but that we are obligated and permitted to not do the trumped obligation.

Chapter Four
Directions for Further Research:
Do We Need a Chimera for Deontic Logic?

There is much intuitive appeal in representing obligation and its cognates in terms of moral necessity. However, this ‘moral necessity’ should not be represented in a straightforward application of classical modal logic. To do so, as we have seen, is to unjustifiably preclude a whole range of moral experiences and widely-supported sorts of moral inquiry: genuine moral dilemmas and trumping-like reasoning. Moreover, classical modal logic is chock full of fallacies of irrelevance, some of which validate deriving an ‘ought’ from an ‘is’.

Thus we have investigated representing obligation and its cognates in terms of non-classical moral necessity. We considered both a relevant modal logic and a sectoring approach to modal logic. While both of these non-classical approaches ably represented fruitful reasoning from inconsistent sets of obligations, we were forced to make revisions to their standard presentations in order to represent trumping-like reasoning. Furthermore, the relevant modal logic straightforwardly precludes the naturalistic fallacy, while again we needed to tinker with Scotch and Jennings’ multiple accessibility relation system so as to jettison this fallacy.

Unfortunately, neither non-classical approach neatly satisfies our proposed criteria. While both systems satisfactorily block obligation explosion when reasoning from an inconsistent set of obligations, substantial revisions were required to account for trumping-like reasoning. However, we notice some similarities in the attempts to use these systems to formalize deontic logic. First, we were justified in postulating that the abstract entities that belonged to the class of morally ideal situations may follow fewer

classical logical laws than do situations like our actual world. Secondly, we were forced to recognize that permissibility does not have the same tidy relationship as obligation does to straightforwardly quantifying over accessible morally ideal situations.

The differences between how we revised both systems to account for trumping-like reasoning lie in part in how readily the points in the relational frames of the corresponding semantics were ripe for such adjustments. While the relevant logic already justified the use of non-logical situations in order to block a long list of fallacies of irrelevance, the sectoring approach of Scotch and Jennings initially used the possible worlds of classical modal logic, albeit with multiple moral accessibility relations. We were unable to simply revise the behavior of these multiple relations so as to represent trumping sufficiently. Non-logical situations were called into action, which prior to this revision were not considered appropriate points in Scotch and Jennings relational frames.

In deontic logic, we notice that the order of premises can be, in fact, relevant in trumping-like reasoning; that is, if \diamond is defined in terms of the truth of formulas in at least one accessible situation. However, in the standard presentation of relevant logic, the order of premises is understood as not being pertinent to a correct characterization of the consequence relation. Thus, we were forced to make changes in the structural rules of relevant logic to accommodate trumping. Moreover, these changes now invalidate the inference $A \rightarrow \sim \sim A$, which is not one of the inferences we were intending to block in appropriate moral inquiry.

In Scotch and Jennings' approach, we were initially able to define \diamond in terms of $\sim \square \sim$, which circumvented the issue of trumped obligations being, nonetheless, permissible. On the other hand, since we were forced to appeal to inconsistent situations,

the classical negation used in Schotch and Jennings' logic will have to be revised, and in so doing, the relationship between \diamond and \square in their system will quickly become complicated, most likely along the lines of that as seen in our discussion of deontic relevant logic.

The use of multiple, possibly disjoint sets of morally ideal situations, via multiple accessibility relations, has much intuitive appeal. We have a robust philosophical interpretation of such semantics, using the notion of multiple cogent ethical theories. Moreover, if we may characterize trumping reasoning as something like a higher-order moral theory which provides reasons why one moral demand should have more weight in our moral inquiry than does another, conflicting demand, then multiple, possibly disjoint sets of morally ideal situations, as picked out by multiple accessibility relations appear, initially, to go far in providing the corresponding semantic structure.

While less dramatic changes seemed to take place in a relevant modal logic in attempts to satisfy our criteria, the philosophical interpretation of its semantics seems less elegant and more complicated—perhaps excessively more—than does the interpretation of Schotch and Jennings' semantics. On the other hand, since cashing out the consequences of revising Schotch and Jennings' semantics with non-logical situations is beyond the scope of this present work, perhaps we would end up with just as cumbersome a philosophical interpretation as it is commonly thought that relevant logic possesses.

In my opinion, to be considered a viable candidate for formalizing deontic logic, a logical system will likely have to use multiple accessibility relations for the class of relevant modal models; a hybrid between a relevant modal logic and Schotch and

Jennings' deontic logic. Moreover, I imagine that such a hybrid's non-logical situations will not only fail to be closed under excluded middle and explosion, but will also fail to be closed under aggregation. While it is merely conjecture, I think that we would then not need to jettison weak commutativity from its corresponding sequent calculus.

Furthermore, if we were able, contrary to how my own intuitions seem to lead, to come up with a palatable semantics for double negation elimination in such a system, then the hybrid's \diamond could have the requisite relationship to \square so as to straightforwardly preclude trumped obligations from being permissible.

Appendix A
Truth conditions for trumping:

We shall write $P < Q$ at x for the situation such that whenever an obligation from type P comes into conflict with an obligation from type Q , the obligation from type Q trumps the obligation from type P , given the information available at x . Likewise, we write $\Box A < \Box B$ at x for the situation such that the set to which $\Box B$ belongs trumps the set to which $\Box A$ belongs, given the information available at x .

Informally, I propose that at $w \models \Box P < \Box Q$, i.e., the claim that the obligation to Q trumps the obligation to P is justified at w if, and only if, for any logical situation x sufficiently similar in the relevant factual and moral respects to w at which P is true, and any situation m of the class of morally ideal situations³¹, there is a logical situation y sufficiently similar in the relevant factual and moral respects to w , at which Q is true, and which is closer to approximating m than x is. The idea is that we say that Q trumps P when, if you were to do some action of type P , you could always have done morally better by doing some action from type Q .

For example, let us assume that I have an obligation to care for a sick parent. Let us also assume, for the sake of argument, that I have an obligation to care for a sick neighbor. However, consider a situation in which the resources to which I have access are so meager, that I can only care for one or the other. Given this assumption, any situation in which both obligations are nonetheless discharged by me will be an impossible situation. We claim that it is true that the obligation to care for a sick parent trumps the obligation to care for a sick neighbor, if and only if, by holding certain relevant facts

³¹ We note that this definition allows that there may not be a unique set of morally ideal situations for any given context. See chapter three for a possible account of formalizing reasoning from multiple, possibly disjoint sets of morally ideal situations.

from the actual situation constant, like a standard upbringing, social milieu etc., across all situations in which I care for a sick neighbor, there will always be a situation, likewise similar in all relevant respects to actuality, in which I care for a sick parent and where, morally speaking, I would do better by so doing.

When we say that, for some p, q where p, q are logical situations, that q is closer than is p to any situation m of the class of morally ideal situations, we mean that fewer relevant factual or moral features would have to be changed at m to exactly resemble the situation q than would be the case for m to exactly resemble the situation p . That is, we might understand ‘ q is closer than p to m ’ as representing that the q situation is a better approximation of m than is the p situation.

Generalizing, it seems reasonable that sometimes there may be three sets or types of situations, let’s use P, Q, R to refer to the class of situation in which $P, Q,$ and $R,$ respectively, are true; it may be that for any situation in which P holds, and for any member of $Q,$ it shares more relevant features with some member of R than any member of $Q.$ For example, consider any situation of type P such that a reasonable but devout leftist would consider it a socialist utopia. Then for any situation (we’ll say it is type Q) in which citizens must commodify their labor in order to be socially productive, there will be a situation of type $R,$ in which production is organized democratically, which would more closely resemble a socialist utopia than would any situation of type $Q.$

This relation between types of situations then suggests that a ternary relation S on sets of situations may be helpful in formalizing trumping.

Definition 5.0 An S relation on a model: given a model $M = \langle sit, logical, R, S, I \rangle$ we say that $SPQR$ if, and only if, for any P -situation x and for any Q -situation y there is a R -situation z which shares more relevant features with x than does y .

Definition 5.1 Truth set: We introduce the definition of a truth set $|A|$ for any formula ‘A’, such that $|A|$ is the set of situations at which ‘A’ is true according to I .

Definition 5.2 The truth set of situations that share all relevant factual and moral features of a situation: We write $|w^*|$ to represent the set of situations which share or make true all the relevant factual and moral features of situation w .

Definition 5.3 The truth set of the class of morally ideal situations: We write $|MI|$ to represent the truth set of the class of morally ideal situations, since we do not preclude the possibility that there may be many morally ideal situations.

Definition 5.4 Trumping in terms of S relation: ‘ $w \vDash P < Q$ ’ is true iff
 $\forall m \forall x \exists y (x, y \in logical) \wedge (x \in |P| \cap |w^*| \wedge y \in |Q| \cap |w^*| \wedge m \in |MI|) \wedge S|MI|PQ$

Proposition 5.0: By the above definition, if $P; Q \vdash f$, then it can never be the case that for any w , $w \vDash P < Q$ and $w \vDash Q < P$ for any P and Q .

Proof: To show this, we suppose that if $w \vDash P < Q$ then:

$\forall m \forall x \exists y (x, y \in logical) \wedge (x \in |P| \cap |w^*| \wedge y \in |Q| \cap |w^*| \wedge m \in |MI|) \wedge S|MI|PQ$

Likewise, suppose also that $w \vDash Q < P$ then:

$\forall m \forall x \exists y (x, y \in logical) \wedge (x \in |Q| \cap |w^*| \wedge y \in |P| \cap |w^*| \wedge m \in |MI|) \wedge S|MI|QP$

However, for any x , if x belongs to the truth set of ‘ P ’, ‘ Q ’ and ‘ w^* ’, then x cannot be logical, contrary to our assumption.

Appendix B

Deontic Negation and Incompatibility

Recall that formulae deduced from the empty set are true at all logical situations, but it is not necessarily the case that such formulae are true at all situations, if we have non-logical situations in our semantics. Since we are considering impossible situations as being possibly morally ideal, it is not valid to infer from formulae that are true at all logical situations that these formulae will also be true at an impossible but morally ideal situation. Therefore, $M; H \vdash f$ does not entail that at non-logical u that $u \models M \rightarrow (H \rightarrow f)$. Thus it is at least an open question whether we may do better by appealing to trumping relations so as to deduce which formulas are negated at accessible morally ideal situations.

Consider the example from chapter two regarding the conflict that an imaginary devout Christian who must decide between giving all his money away and caring for a sick dependent. We also stipulated that caring for your dependents should over-ride giving all your money away:

$M; H \vdash f$ and $s \models M < H$ but $s \not\models \Box M$ and $s \not\models \Box H$, where M stands for ‘giving all your possessions to the poor’ and H stands for ‘caring for your sick child’.

What we would like to be able to infer is that it is obligatory and permissible to not give away all your money, given that you have a sick child to care for. We do not want it to be permissible nor obligatory to not care for your child.

Consider a logical s such that $s \models M < H$ and $s \models \Box M \& \Box H$ but that $M; H \vdash f$. We now introduce a new incompatibility relation N° for deontic logic such that:

Definition 6.0 With a model $\mathbf{M} = \langle sit, logical, S, R, N^\circ, I \rangle$, N° is a binary relation on sit such that $N^\circ xy$ iff for any P -situation y there is an Q -situation x such that $S \mid MI \mid PQ$ and for any x of $|Q|$ and for any y of $|P|$, $P; Q \vdash f$ or $Q; P \vdash f$ (for any logical situation)

Definition 6.1 \sim negation in terms of N° :

for any $m \in |MI|$, ' $m \vDash \sim A$ ' iff for any y , there is an x such that if $y \vDash A$ then $N^\circ xy$

We can't have the relation N° being symmetrical, thus we also have to reject weak commutativity as a structural rule in any corresponding sequent calculus. And by adding $\leftarrow I$ as a rule, we add a new negation operator, similar to chapter two.

Definition 6.2 \neg -negation in terms of N° :

for any $m \in |MI|$, ' $m \vDash \neg A$ ' iff for any y there is no x such that if $x \vDash A$ then $N^\circ xy$

Suppose that at s , for any logical M situation M an agent would always do better by realizing some logical H situation H instead, given that M and H are sufficiently similar to s .

Therefore, since we can represent this as:

$$s \vDash \forall m \forall x \exists y (x, y \in \text{logical}) \wedge (x \in |M| \cap |s^*| \wedge y \in |H| \cap |s^*| \wedge m \in |MI|) \wedge S|MI|MH$$

we can infer that $s \vDash M < H$

Claim With the incompatibility relation N° we have semantic tools for using the hierarchy function in chapter one. We show that:

Since $s \vDash M < H$, we infer that for any m of the class of morally ideal situations, that $S|MI|MH$. Thus we have for any situation at which H holds and for any situation at which M holds, that $N^\circ HM$ which according to our conditions above, gives us:

- 1) $s \vDash \Box H \rightarrow \Box \sim M$
- 2) $s \vDash \Box M \rightarrow \Box \neg H$ ³²

and these formula are the same as the results we desire from our hierarchy.

Recall that the h hierarchy function from chapter one:

$$s \vDash \Box M \ \& \ \Box H \text{ and } s \vDash h: \text{ for any } M, H \text{ such that } [((\Box H \ \& \ \Box M) \text{ but } M; H \vdash f) \ \& \ \text{for some } P, Q \text{ } ((\Box M \in P \ \& \ \Box H \in Q) \vee (\Box H \in Q \ \& \ \Box M \in P)) \ \& \ \sim((\Box M \in P \ \& \ \Box H \in P) \vee (\Box H \in Q \ \& \ \Box M \in Q))] \ \& \ (P < Q \vee Q < P \ \& \ \sim(P < Q \ \& \ Q < P))^{33}] \rightarrow [((Q < P \ \& \ \Box M \in P) \rightarrow \Box \sim H) \vee ((Q < P \ \& \ \Box H \in P) \rightarrow \Box \sim M) \vee ((P < Q \ \& \ \Box M \in Q) \rightarrow \Box \sim H) \vee ((P < Q \ \& \ \Box H \in Q) \rightarrow \Box \sim M)]$$

Proof: We claim that if $s \vDash \Box M \ \& \ \Box H$ and $s \vDash M < H$ then $s \vDash \Box \sim M \ \& \ \Box H \ \& \ \Box M \ \& \ \Box \neg H$ which is what we desire from our hierarchy. We have that for any x if Msx then $x \vDash M$ and $x \vDash H$. However, since for any w of s if w is logical, then $w \vDash (M \ \& \ H) \rightarrow f$, therefore $x \notin \text{logical}$. From $s \vDash M < H$ we infer that for any x such that Msx then, since $s \vDash \Box M \ \& \ \Box H$, we have that $x \in |MI|$ and thus $S|MI|MH$ and therefore $N^\circ HM$. So: for any x such that Msx , $x \vDash \sim M \ \& \ \neg H$ by both negation definitions. Since for any x such that Msx , x is

³² We can interpret these as follows: 1) if you are obligated to care for your sick child, you are obligated to not give all your money away; 2) if you are obligated to give all your money away, then it is impermissible to do actions which preclude also caring for your sick child.

an inconsistent situation, we have $x \vDash M \& H \& \sim M \& \neg H$. Therefore we have as the desired result: $s \vDash \Box \sim M \& \Box H \& \Box M \& \Box \neg H$.

According to the above **h**, if $M; H \vdash f$ and $s \vDash \Box M \& \Box H$ and $s \vDash P < Q$ and $\Box H \in Q$ then **h**: $\Box \sim M$. There is no model, under these assumptions, that satisfies both $N^\circ QP$ and $\Box \sim H$, since we have both negations defined in terms of a non-symmetrical N° for morally ideal situations; To show this consider that if $s \vDash P < Q$ then, under our assumptions, for any x such that Msx , then x cannot $\vDash \sim H$, since it is never the case that $N^\circ MH$ as shown in the corollary in appendix A.

Q.E.D.

Works Cited

- Bell, John et.al. Logical Options. Peterborough: Broadview Press, 2001.
- Delong, Howard. A Profile of Mathematical Logic. New York: Dover Publications, 2004.
- Dewey, John. Essays in Experimental Logic. New York: Dover Publications, 2004.
- Chellas, Brian F., King-Farlow, John, and Nielsen, Kai et al. Canadian Journal of Philosophy, Volume VII, Number 2. "Moral Dilemmas and Consistency in Ethics", McConnell, Terrance C. Canadian Association for Publication in Philosophy, 1978.
- Gabbay, Dov and Franz Guentner. Handbook of Philosophical Logic Volume 4. "Logics for Defeasible Argumentation", Prakken, H., Vreeswijk, G. Boston: Kluwer Academic Publishers, 2001.
- Gabbay, Dov and Franz Guentner. Handbook of Philosophical Logic Volume 6. "Relevance Logic", Dunn, Mike, Restall, Greg. Boston: Kluwer Academic Publishers, 2001.
- Gabbay, Dov and Franz Guentner. Handbook of Philosophical Logic Volume 6. "Paraconsistent Logic", Priest, Graham. Boston: Kluwer Academic Publishers, 2001.
- Gabbay, Dov and Franz Guentner. Handbook of Philosophical Logic Volume 8. "Deontic Logic", Åqvist, Lennart. Boston: Kluwer Academic Publishers, 2001.
- Gabbay, Dov and Franz Guentner. Handbook of Philosophical Logic Volume 12. "How to go Monotonic", Mackinson, David. Boston: Kluwer Academic Publishers, 2001.
- Goodman, Nelson. Fact, Fiction, and Forecast. Cambridge: Harvard University Press, 1983.
- Gowans, Christopher. Moral Disagreements. "Moral Disagreement", Brink, David. New York: Routledge, 2000.
- Hilpinen, Risto. New Studies in Deontic Logic: Norms, Actions, and the Foundations of Ethics Volume 152. "Non-Kripkean Deontic Logic", Schotch, Peter, Jennings, Raymond E. Berlin: Springer, 1981.
- Häyry, Matti and Tuija Takala. The Future of Value Inquiry Volume 112. "Two Types of Moral Dilemma", Welch, John R. City: Editions Rodopi, B.V, 2001.
- Mackie, J. Ethics. Harmondsworth Eng.: Penguin, 1977.

Mares, Edwin. *Relevant Logic*. Cambridge Etc.: Cambridge University Press, 2004.
Works Cited

Nute, Donald. *Defeasible Deontic Logic* Volume 263. "Defeasible Reasoning and Moral Dilemmas", Koons, Robert C., Seung, T.K. Boston: Kluwer Academic Publishers, 1997.

Restall, Greg. *An Introduction to Substructural Logics*. New York: Routledge, 2000.